RĂZVAN RUGHINIȘ

STĂPÂNI, SOCIETATEA CETĂȚENI SAU DIGITALĂ SCLAVI?

ebook HUMANITAS

Răzvan Rughiniș, profesor la Facultatea de Automatică și Calculatoare a Universității Politehnice din București, doctor în știința calculatoarelor din 2005 și, din 2021, absolventul programului de master „Etică aplicată în societate, afaceri și organizații“ la Facultatea de Filozofie a Universității din București, urmărește crearea de comunități și de traiectorii de termen lung în educație și transformare socială prin tehnologie. În 2013, alături de antreprenorul și investitorul Andrei Pitiș, a pus bazele programului Innovation Labs (www.innovationlabs.ro), care a sprijinit peste 500 de noi produse digitale prin start-upuri și spin-offuri, având o amprentă națională în 18 centre universitare din România. Cercetarea și cursurile sale se axează pe securitate cibernetică, internetul lucrurilor și protecția vieții private. În ultimii 10 ani, a publicat, în calitate de autor și coautor, peste 50 de lucrări științifice pe tema securității și protecției datelor personale. De asemenea, a publicat peste 30 de articole științifice în domeniul interacțiunii om–calculator, cu accent pe studiul experienței utilizatorilor, tehnologii persuasive și gamificare, inclusiv în conferințe și jurnale prestigioase ale domeniului precum ACM-CHI, ACM-AltChi, CSCL și HCI, *Computers & Security*, *Computer Law & Security Review* și *Ethics and Information Technology*. Este membru al Academiei Oamenilor de Știință din România.

RĂZVAN RUGHINIȘ

SOCIETATEA DIGITALĂ

STĂPÂNI, CETĂȚENI SAU SCLAVI?

Cuvânt înainte
de Constantin Vică

Acest e-book este protejat de legea drepturilor de autor. Reproducerea sa integrală sau parţială, multiplicarea sa prin orice mijloace şi sub orice formă, punerea sa la dispoziţie publică pe internet sau în reţele de calculatoare, stocarea sa temporară sau permanentă pe diverse dispozitive sau în sisteme care permit recuperarea informaţiei, gratuit sau în scop comercial, precum şi alte fapte similare, fără permisiunea scrisă a editurii reprezintă o încălcare a legislaţiei privind protecţia dreptului de autor şi se pedepsesc conform legilor în vigoare.

© HUMANITAS, 2023 (ediţia digitală)

ISBN: 978-973-50-7804-1 (epub)

EDITURA HUMANITAS
Piaţa Presei Libere 1, 013701 Bucureşti, România
tel. 021.408.83.50, fax 021.408.83.51
www.humanitas.ro

Comenzi online: www.libhumanitas.ro
Comenzi prin e-mail: vanzari@libhumanitas.ro
Comenzi telefonice: 0723.684.194

Mamei mele, Domnica, pentru care știința a fost mereu
un instrument în slujba înțelepciunii

Cuprins

Cuvânt înainte
Introducere
1. Tehnologii digitale, capitalism și modele sociale: SUA, UE și China
    Autonomia sferelor justiției
    Regimuri ale inegalității
    Deliberarea în sfera publică
2. Preaplinul cunoașterii și minunata lume nouă a puterii algoritmice
    Preaplinul cunoașterii: transformarea informației în valoare adăugată
    Cunoașterea intimă: relevanța și personalizarea
    Cunoașterea extinsă: expansiunea asocierilor, grupurilor și rețelelor
    Cunoașterea autonomă și pseudoumană: inteligența artificială
    Cunoașterea-combustibil: de la informație, la accelerație
3. Excesul cunoașterii în societatea algoritmică
    Excesul de precizie: dezanonimizarea datelor și daunele voluntare sau involuntare asupra indivizilor
    Dimensiuni ale excesului de precizie
    Exces de precizie, sau mistificarea unor coincidențe?
    Exces de accesibilitate și dreptul de a fi uitat
    Exces de informații ilegitime
    Excesul costurilor cunoașterii
    Excesul relativ: (dez)echilibrul cunoașterii
4. Deficitul cunoașterii în societatea algoritmică
    Deficitul de precizie: erori ale cunoașterii bazate pe Big Data
    Erorile aleatorii
    Erorile sistematice
    Deficitul de înțelegere: iluzia cunoașterii
    Cauzalitate și corelație
    Cutiile negre din deciziile algoritmice, din învățarea automată și din inteligența artificială
    Deficitul de înțelepciune: dezumanizarea indivizilor ca data points
    Deficitul de problematizare etică a științei datelor
    Deficitul de adevăr factual: Big Data, știrile false și infodemia
    Agnosticismul algoritmilor de recomandare

Golurile de date și prăbușirea în lumi contrafactuale
Deficiențe adversariale: războiul cunoașterii în societatea algoritmică
5. În balanță: noi riscuri morale ale științei și societății algoritmice
6. Regulile jocului: spre un model european al autonomiei umane
Etica și drepturile omului în societatea algoritmică: filozofia morală și modelul european
De la vorbe la fapte: utilizatori, experți, cetățeni
Utilizatori și cumpărători
Specialiști și profesioniști
Cetățeni și decidenți
Mai mult, mai repede, mai opac: noi fronturi ale inteligenței artificiale
Pandemia COVID-19 și legitimarea IA prin managementul algoritmic
Criminalizarea avortului și legitimarea IA prin supraveghere algoritmică
Ucraina și legitimarea IA prin armata algoritmică
Contraofensiva interesului public
De la societatea cunoașterii la capitalismul supravegherii
Mulțumiri
Creditele tabelului și figurilor
Note

Cuvânt înainte

Răzvan Rughiniș este unul dintre adolescenții minunați ai revoluției digitale și internetului. Fără a fi un nativ digital, profesorul de la Facultatea de Automatică și Calculatoare a Politehnicii bucureștene a crescut odată cu internetul, dezvoltând, la rândul său, rețeaua rețelelor, într-o simbioză conceptuală care astăzi, din păcate, este greu posibilă. Este printre puținii cercetători români care investighează, în deplină cunoaștere a tehnologiei rețelelor, rolul ei social și impactul moral pe care îl are asupra vieții noastre. Ca profesor, a ajutat și ajută studenții, generație după generație, să descopere ce se poate construi prin creația computațională și designul rețelelor, la cursuri precum Introducere în sisteme de operare și Tehnologii de protecție a vieții private, și, mai ales, cum să vadă dincolo de actul ingineresc în sine semnificațiile sale. Ca activist pentru drepturile utilizatorilor online, Răzvan este, în momentele critice, la masa negocierilor, online sau în stradă, un apărător fervent al libertății internetului și al libertăților noastre digitale. Cercetătorul susține, de ani de zile, o întreagă agendă, integrativă, a acestor probleme, fiind în avangardă. Dacă am căuta o formulă să-l descrie, cred că aceasta i s-ar potrivi cel mai bine: Răzvan Rughiniș este avangardistul critic și rafinat al unei tehnologii în continuă expansiune. Progresul pe care îl urmărește nu este doar științific și tehnologic, așa cum se întâmplă în majoritatea cazurilor în lumea tehnologiei informației, ci, în primul rând, social și moral. Cartea pe care o aveți în mâini acum este cea mai recentă și clară mărturie.

Am avut norocul de a lucra împreună în perioada în care a fost student al masterului de etică aplicată de la Facultatea de Filozofie. Profesorul de informatică și internet nu s-a sfiit să fie și studentul teoriilor morale și politice. Atunci am văzut cum începe această aventură intelectuală, pe care acum o putem citi împreună, online sau offline, pe hârtie sau digital.

Lumea digitală de astăzi, mai ales la final de pandemie, mai păstrează prea puțin din promisiunile democratice ale internetului și *world wide web*-

ului din stadiile lor incipiente, din perioada de aur, nereglementată, a primilor exploratori online, perioadă care s-a încheiat odată cu trecerea la web 2.0, platforme și ecosisteme totale, așa cum sunt Google, Facebook sau Apple. Libertatea de comunicare și expresie scade de la zi la zi, arată toate rapoartele organizațiilor internaționale, autonomia cognitivă și cea morală sunt reduse sub fluxul îmbătător al informației, lucru care ni se întâmplă tuturor, iar despre egalitatea vocilor și recunoașterea cu drepturi depline aproape că nimeni nu mai vorbește, chiar dacă nedreptățile, de la cele epistemice, cum ar fi dezinformarea, la discriminare și obiectificare, se țin lanț. A devenit o lume a captologiei și dependenței structurale, a dataficării, a transformării vieții în date și a procesării lor algoritmice. Și în tot acest vortex al noii cunoașteri, în care *Big Data* și inteligența artificială sau învățarea automată devin instrumente cuantificatoare ale puterii, Răzvan vrea să facă lumină, dând dovada forței sale intelectuale și a spiritului său critic. Este, într-un fel, un alt răspuns, diferit, dar nu întru totul, de cel kantian, la întrebarea „Ce este luminarea?“ Pentru că răspunsul său nu este o mărturisire despre viața online, nici un manual de supraviețuire în societatea digitală, ci chiar uzul public al rațiunii împotriva raționalizării numerice. Ca un învățat, cum ar fi spus Kant, își urmează datoria morală publică de a denunța pericolele și capcanele abstractizării algoritmice, care se hrănește din date fără ipoteze și din modele fără viață, chiar dacă această datorie ar putea intra în conflict cu cele de profesor-inginer și promotor al inovației digitale.

Conflictul datoriilor nu are loc, pentru că propunerea cărții nu este o chemare la revoluție (un termen uzurpat de retorica salvaționismului tehnologic, retorică promovată de marile companii, dar și de state autocrate, precum China sau chiar democratice, precum SUA), ci la transformarea rafinată a raporturilor pe care le avem cu giganții tehnologici ai lumii online. Mesajul pe care cartea îl transmite este optimist și, astfel, constructiv: lupta nu e pierdută, o altă așezare, chiar obștească, născută din voința politică a celor care trăiesc postdigital – distincția digital/non-digital nu-și mai are sensul astăzi – este posibilă. Tehnologiile digitale joacă rolul noilor instituții, impun cadrele normative sociale și extind sau restrâng posibilități de acțiune, coordonare sau colaborare, dar ele nu plutesc într-un

vacuum. Platformele social media, pieţele cu două laturi la masa cărora stăm online, şi nu ne gândim doar la Amazon, nu au o forţă destinală, ci sunt contingente, ţin de contextul mai larg al instituţiilor politice şi sociale ale societăţilor în care se dezvoltă. Chiar dacă utilizarea acestor platforme, aplicaţii şi servicii online ne in-formează, ne de-formează şi ne transformă, la rândul lor ele pot fi îndreptate şi orientate în acord cu valorile noastre morale exprimate prin instituţii. Constructivismul social e alegerea corectă, determinismul e doar o altă sirenă din corul retoric al salvării prin progres tehnologic nemăsurat.

Aproape toţi, până şi copiii, vorbim astăzi despre Big Data (date vaste) şi inteligenţa artificială (sisteme produse mai ales prin învăţarea automată, *machine learning*), despre supercomputere, unele cuantice, şi internetul lucrurilor. Dar prea puţini privim dincolo de interfaţa aproape psihedelică a maşinăriilor care ne mediază existenţa. Răzvan o face filozofic aplicat, dincolo de interogativ şi deloc moralist de paradă, având capacitatea ştiinţifică şi inginerească a celui care ştie să folosească linia de comandă spre a controla maşina.

Cartea pe care o răsfoiţi analogic sau digital este un manifest pentru un model politic european de integrare justă şi dreaptă a proceselor algoritmice, bazate pe date vaste, în ţesătura socială, integrare care să aducă nu doar profit companiilor, ci o bunăstare necesară tuturor şi fiecăruia în parte. Felul în care ajunge să argumenteze pentru un model european, care e capabil să aşeze respectul pentru persoane, pentru autonomia lor, şi non-discriminarea înaintea capitalismului supravegherii, este provocator. Lumea e mai transparentă odată cu posibilitatea de a-i capta aproape toate fenomenele şi de a le reprezenta numeric. Avem cunoaştere din belşug, care se poate (re)produce nelimitat, căci fluxul datelor ţine ritmul fluxului vieţii. Se poate şti despre fiecare dintre noi mai mult decât am fi aflat vreodată prin introspecţie. Dar nu neapărat această cunoaştere, de ordin pozitiv, reuşeşte să surprindă ce suntem şi ce ne dorim cu adevărat. De altfel, comportamentele noastre sunt atât de bine descifrate, încât sub pecetea tainei a mai rămas doar speranţa că nu ne-am transformat în zombi sau borgi. Cunoaşterea prin date vaste, cu preaplinul şi excesele sale, cu deficitul său care provine din implicita lăsare deoparte a lumii vieţii aşa

cum e trăită, subiectiv, de fiecare minte umană în corpul ei, nu e în stare să problematizeze. Cititorii de filozofie românească pot auzi un ecou din încercarea lui Lucian Blaga de a dezvălui procesul cunoașterii luciferice, cea înțelegătoare. Deficitul de înțelegere, ne spune și Răzvan Rughiniș, este produsul unui proces de cunoaștere care nu servește dezlegării nici unui mister, care elimină și caracterul concret, intuitiv, al experienței, dar care proces are rolul de a guverna tranzacțional persoanele și bunurile lumii postdigitale, astfel încât să se potrivească, într-o simfonie a plăcerii și fără nici o reacție critică. Agnosticismul algoritmilor în ceea ce ne privește, nepăsarea lor față de convingerile umanității, dar și faptul că au un caracter inanalizabil (fiind opaci investigației umane, chiar și privirii tehnice a specialiștilor) îi plasează în categoria secretului. Îndeajuns de ironic, dacă înțelegem că noi suntem aproape nuzi în fața lor, a companiilor și a statelor care-i controlează, în timp ce ei ne controlează dintr-o zonă de obscuritate sau chiar de invizibilitate.

Într-o economie a atenției, guvernată prin algoritmi și condusă de date, într-un sistem al supravegherii continue și al satisfacerii prompte, dar nesatisfăcătoare, a oricărei dorințe, într-o lume cândva unită de internet, acum divizată (și) de aplicațiile sale, cartea lui Răzvan Rughiniș este o călăuză abilă, care ne face să avem ochii deschiși când traversăm împreună jungla deșertăciunii online, prin travaliu moral, spre regăsirea autonomiei.

Constantin Vică
București, 2022

Introducere

Pentru unii mumă, pentru alţii ciumă.

Trăim într-o societate care glorifică informaţia şi tehnologia. Animaţi de câştigurile şi promisiunile noii lumi digitale, iertăm uşor şi uităm rapid limitele şi victimele acesteia. Avansul tehnologiilor de modelare a realităţii prin Big Data şi de acţiune pe baza deciziilor algoritmice a generat un nou val al automatizării stimulate de inteligenţa artificială. Prinşi în goana inovării, riscăm să neglijăm întrebări cruciale despre noua societate pe care o fabricăm. Cine va controla cunoaşterea generată prin extinderea realităţii în mediul digital? Cine va împărţi beneficiile acestui tip de cunoaştere şi cui anume? Cine va sesiza erorile şi câte dintre ele vor fi vizibile? Cine va fi responsabil pentru decizii şi cine va plăti pentru consecinţele lor?

Într-o seară de duminică, în mai 2020, Michael Williams, un bărbat de 65 de ani din Chicago, a ieşit cu maşina să-şi cumpere ţigări de la benzinărie. Era o seară agitată; oraşul era cuprins de protestele violente declanşate de uciderea lui George Floyd de către poliţie, cu şase zile înainte. Conform relatării lui Williams, povestită de *Associated Press*[1], benzinăria fusese vandalizată. Bărbatul a pornit înapoi spre casă. Pe drum, un tânăr pe care îl cunoştea, din cartier, i-a făcut cu mâna, să-l ia şi pe el cu maşina. Williams s-a oprit şi Safarian Herring, de 25 de ani, s-a urcat pe scaunul din dreapta. Seara a luat o turnură catastrofică: la o intersecţie, o maşină s-a oprit lângă ei şi şoferul necunoscut l-a împuşcat pe Herring în cap. Williams nu a fost atins de glonţ. Şocat şi îngrijorat pentru soarta pasagerului care sângera vizibil, l-a dus la urgenţă. Câteva zile mai târziu, tânărul a murit. Câteva luni mai târziu, în august, Williams a fost arestat şi pus sub acuzare pentru crimă, fiind principalul suspect în uciderea lui Herring. Dovezile împotriva lui nu proveneau însă de la martori oculari sau analize balistice, ci de la sistemul de inteligenţă artificială ShotSpotter, menit să detecteze focurile de armă în spaţiul public. Sistemul detectase un zgomot care, pe parcursul investigaţiei, a fost diagnosticat ca reprezentând un foc de armă tras din

maşina lui Williams. Pe baza acestor informaţii, Williams a fost ţinut în închisoare timp de 11 luni. El nu a ajuns să fie judecat, deoarece, după aproape un an petrecut în detenţie, procurorii au renunţat la caz, considerând că nu au suficiente dovezi. Williams a fost eliberat. Nu este deloc clar cine va plăti pentru această greşeală, care l-a costat pe Williams 11 luni de coşmar – sau dacă va plăti cineva. Oraşul Chicago este cel mai mare client al ShotSpotter, având un contract încheiat pe trei ani, în valoare de 33 de milioane de dolari. Speranţele municipalităţii sunt să folosească astfel de soluţii de inteligenţă artificială pentru a identifica ariile de risc şi a controla mai bine problemele de ordine publică. Totuşi, o evaluare independentă a ShotSpotter, publicată în 2021[2], a ajuns la concluzia că nu există vreun beneficiu al folosirii acestui produs în ceea ce priveşte reducerea mortalităţii provocate de focuri de armă sau îmbunătăţirea ratelor de arestare. În ce măsură inteligenţa artificială contribuie, în acest caz, la buna cunoaştere şi justa decizie în păstrarea ordinii publice, sau, dimpotrivă, conduce la noi surse de eroare şi arbitrar?

Ai carte, ai parte, spune înţelepciunea secolelor trecute. Atât înainte, cât şi acum, deţinerea de cunoştinţe la nivel personal era şi este considerată o valoare. Astfel, cunoaşterea personală se caracterizează prin conversia recurentă, la o rată de schimb destul de bună, într-un plus de autonomie şi bunăstare. Actualmente însă, cunoaşterea nu se mai află doar în clasicele biblioteci pline cu cărţi, ci şi în rezervoare informaţionale digitale. Noile depozite de informaţii – stocate în *cloud* şi procesate prin învăţare automată şi inteligenţă artificială – au în vedere atât vastul internet, unde orice informaţie poate fi accesată prin motoare de căutare, cât şi platformele colaborative, Wikipedia, reţelele de socializare şi Big Data, ca să enumăr doar câteva dintre bibliotecile moderne. Această expansiune a cunoaşterii, din mediul fizic, palpabil, în mediul digital, virtual, schimbă regulile jocului, înlocuind atât cărţile din mână, cât şi cele din mâneca jucătorilor. Alături de realizări nesperate şi promisiuni înfierbântate, societatea informaţională aduce şi noi derapaje şi riscuri morale.

Cunoaşterea este o monedă cu două feţe. Pe de o parte, oferă emancipare şi deschide orizonturi. Pe de altă parte, sub forma sa informaţională, permite controlul nu doar asupra ei înseşi, ci şi asupra mediului extern,

inclusiv asupra celorlalţi. Cum vor evolua carierele şi meseriile unor generaţii umane intrate deja în competiţie directă, pe piaţa muncii, cu boţi şi procese robotice? Ce se întâmplă când sfera privată ne este străpunsă de organizaţii comerciale, politice sau statale prin tehnologii digitale ce recoltează continuu şi dau mai departe, pe faţă sau pe ascuns, informaţii minuţioase despre gesturile şi preferinţele noastre? Cum se reconfigurează raportul de valoare dintre cunoaşterea pe care o câştigăm şi informaţia care se scurge din noi?

Ca profesor şi inginer în domeniul calculatoarelor şi tehnologiei informaţiei, sunt în permanenţă confruntat cu transformarea tot mai rapidă a tehnologiilor şi modelelor de afaceri asociate cu inovaţiile digitale. În 2013, alături de Andrei Pitiş, am pornit InnovationLabs[1], un accelerator pentru start-upuri de tehnologie, în care participă anual tinere echipe din universităţile româneşti. Mă regăsesc, aşadar, într-o triplă ipostază: de inginer, de mentor în ecosistemul de start-upuri ICT şi de profesor universitar. Fiecare dintre aceste roluri au afinităţi diferite cu noile tehnologii digitale. Ca inginer, împărtăşesc, alături de studenţii şi absolvenţii facultăţilor de calculatoare şi de informatică, pasiunea pentru frumuseţea intrinsecă a tehnologiilor digitale şi pentru puterea lor transformativă. Ca mentor într-un program pentru tineri antreprenori, sunt preocupat de diferenţa dintre tehnologii şi produse tehnologice. Nu tehnologiile abstracte sunt cele care schimbă lumea, ci felul în care sunt transpuse în produse reale. Acestea ajung să aibă impact prin lansarea pe piaţă, cel mai adesea, conform unei strategii calculate pentru expansiune şi profit. Produsele ce operează integral într-un model nonprofit, precum Wikipedia sau OpenStreetMap, sunt relativ rare în economia internetului. Ca profesor şi cercetător universitar, sunt implicat în dialogul ştiinţific şi studiile critice privind impactul social şi mizele etice ale acestor produse. Pe lângă entuziasmul ingineresc şi optimismul specific mediului antreprenorial, am ajuns, astfel, în contact cu scepticismul şi, uneori, indignarea sau furia celor ce observă erorile, riscurile şi abuzurile intermediate de noile tehnologii digitale şi modelele lor de afaceri.

Inteligenţa artificială, bazată pe Big Data, restructurează radical şi economia, şi ordinea politică şi socială. În România, după o revoluţie şi o

tranziție, iată-ne acum antrenați într-o accelerare colectivă spre o destinație incertă, alături de concetățenii europeni și de conviețuitorii planetei Pământ.

Concepte fundamentale, precum libertate, justiție, decizie, inteligență, proprietate, intimitate, adevăr, prieten sau comunitate, își schimbă sensurile gradual, aproape imperceptibil, de la o zi la alta și riscă să sufere redefiniri radicale. Avem deja semne clare ale importanței deciziilor algoritmice, datelor vaste și tehnologiilor asociate în viața noastră de zi cu zi. Printre acestea se numără controversele recente privind rolul pe care giganticele corporații de tehnologie, așa-numitele Big Tech (Google, Facebook, Amazon, Apple, Microsoft etc.), îl au în diverse situații. Astfel, destructurarea democrației și creșterea excluziunii, precarității și inegalității sociale sunt legate de geneza unui nou tip de capitalism, al supravegherii (*surveillance capitalism*). O serie de investigații jurnalistice și studii științifice privind erori și discriminări algoritmice – o parte dintre ele discutate în capitolele de mai jos – surprind derapaje frecvente de la traiectoriile promise. Acestea nu anulează potențialul lor de creare de valoare socială și economică, dar solicită o examinare atentă, chiar și din perspectivă etică, a costurilor și distribuției acestei valori, pe lângă punctele de vedere inginerești sau economice.

Puterea economică a companiilor Big Tech a crescut substanțial în anii de pandemie, acestea ajungând la începutul anului 2022 să fie mai valoroase financiar decât economiile unor state din vârful ierarhiei globale. La aproape trei miliarde de dolari, Apple a depășit în decembrie 2021 valorile anuale ale economiilor Marii Britanii, Indiei sau Franței. Microsoft, cu o capitalizare pe piață de aproximativ două miliarde și jumătate de dolari, a depășit Rusia, Canada sau Coreea de Sud.[2] Această evoluție este problematică, deoarece subliniază fragilitatea buclelor de reglaj din societățile capitalismului datelor.

Auzim adesea afirmări ale valorii simplității. Pentru cetățenii unei lumi încâlcite și accelerate, minimalismul stilului de viață sau puritatea unui stil artistic rezonează puternic emoțional. În același timp, trăim vremuri în care valoarea unei persoane ține de specificul său de a fi în lume, de nuanțele pe care le aduce mediilor statistice prin propria sa gândire și experiență. Solicităm să fim văzuți ca indivizi, nu ca numere sau ca elemente ale unui

colectiv omogen. Cerem același lucru și pentru concetățenii noștri. *Rezultă de aici o societate tot mai complicată, dar în care complexitatea este bună*, funcționând ca un garant al puterii individuale de expresie. Reversul medaliei pentru această complexitate pozitivă constă în creșterea riscului de erori, de procesare inadecvată a cazurilor particulare și atipice, de anticipare greșită a efectelor unor decizii sau de marginalizare a celor lăsați în urmă de noi limbaje și tehnologii. Societățile democratice se bazează pe bucle de reglaj la toate nivelurile de decizie, încercând să prevină și să remedieze astfel de derapaje. Vorbim, așadar, despre asigurarea transparenței decizionale, astfel încât publicul să poată observa când deciziile sunt bazate pe argumente și calcule greșite. Mai mult, prin drepturi de liberă exprimare și drepturi de vot, membrii comunităților pot să îi sancționeze pe liderii care și-au asumat pariuri neinspirate sau păguboase. Toate aceste procese de transparentizare și de feedback se aplică diferit în cazul corporațiilor.

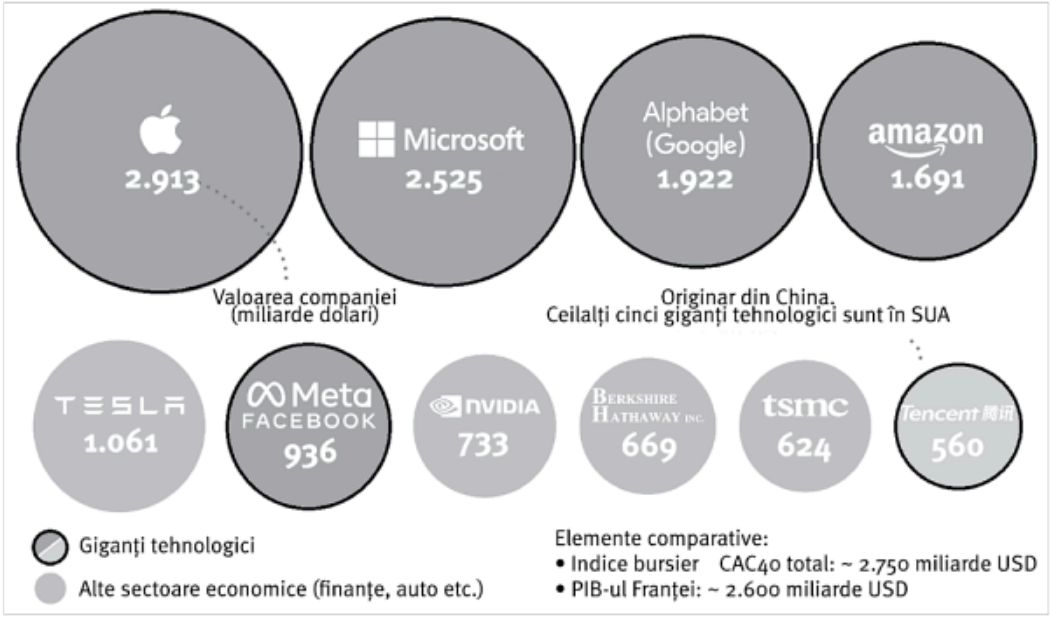


Fig. 1 Cele mai mari zece corporații în funcție de capitalizare, la finalul anului 2021, comparate cu indicele bursier CAC40 și cu PIB-ul Franței. Sursă: YBSLE, imagine publicată pe Wikimedia Commons.

Spre deosebire de logica democratică, orientată spre interesele diverse ale cetățenilor, marile corporații trebuie să dea cu adevărat socoteală doar acționarilor. Cu cât sunt mai mari, cu atât ele câștigă puterea de a ignora

cerințele angajaților sau ale statelor în care operează. Buclele de reglaj ale corporațiilor sunt orientate spre metrici pe care publicul, adesea, nu le vede, urmărind, ca prioritate centrală, maximizarea profitului. Mai mult, temporalitatea deciziilor și adaptărilor corporațiilor iese din scala dezbaterilor democratice. La intersecția accelerării tehnologice și a dinamicii piețelor globale, marile organizații comerciale joacă un joc în care drepturile și valoarea vieții indivizilor din societățile în care acționează apar ca niște numere optimizabile în foi de calcul cost–beneficiu, nu ca valori fundamentale.

O putere economică vastă se traduce și în putere politică vastă. Companiile își promovează interesele proprii prin finanțarea, cu bugete considerabile, a intervențiilor de tip lobby, a candidaților în campanii electorale, a avocaților care le apără acțiunile în justiție și a studiilor care le oferă legitimitate.

Companiile Big Tech se remarcă nu doar prin mărimea absolută și puterea lor financiară, ci și prin ritmul rapid de creștere, ce depășește cu mult dinamica economiilor naționale. În ultimii ani, creșterile anuale ale acestor corporații au depășit adesea 20%, în contextul unor creșteri de maximum 5% ale economiilor naționale (vezi figura 2). Prin urmare, se petrece o sporire continuă a puterii unor companii ce-și urmăresc propriul profit în spatele unei cortine opace create din algoritmi indescifrabili ai inteligenței artificiale, cantități inimaginabile de date preluate de la utilizatori și invocări ale secretului comercial.

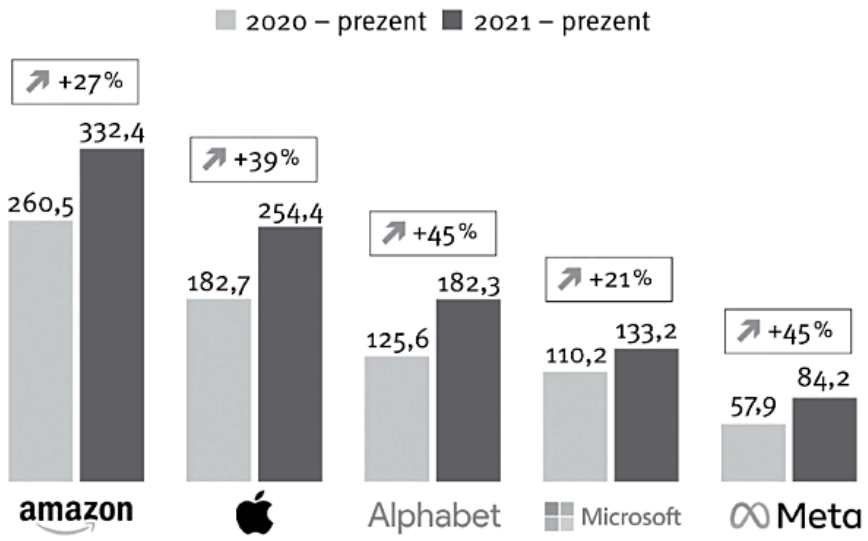


Fig. 2 Ritmul creșterii marilor giganți tehnologici în perioada pandemiei. Sursă: Statista.

Societatea bazată pe infrastructurile marilor companii de tehnologie, pe inteligență artificială și pe colectarea continuă de date despre indivizi este, cel mai adesea, descrisă fie *utopic*, fie *distopic*. În scenariile optimiste, tehnologiile digitale unesc omenirea și deciziile algoritmice se ridică deasupra îngustimii minții umane, depășind interesele egoiste, calculele mioape, părtinirile și erorile de raționament. Prin anvergura considerațiilor și precizia estimărilor, procesele algoritmice reușesc, în cele din urmă, să atingă beneficiile promise. Astfel de rezultate favorabile vizează salvarea umanității pe termen lung, eliminarea discriminării, implementarea unei meritocrații perfecte și adaptarea continuă la provocări.

Prin contrast, în scenariile pesimiste, povestea de mai sus este doar versiunea ideologică pe care un Frate cel Mare orwellian, fie corporativ sau politic, o servește populațiilor controlate prin supraveghere continuă și decizii infinit personalizate. Realitatea societății algoritmice este descrisă, distopic, prin aservirea muncitorilor umani unor șefi robotici ce-i elimină prompt pe cei inadecvați, prin controlul total al fluxurilor informaționale și prin manipularea maselor, trăgând de ițele fine ale preferințelor și vulnerabilităților individuale, divulgate continuu pe platforme și dispozitive digitale. Privind prin această lentilă, vedem cum profilarea digitală, clasificarea și deciziile algoritmice nu servesc binelui omenirii, ci intereselor elitei înguste care construiește și controlează aceste procese,

extrăgând cea mai mare parte din valoarea creată. În loc de un *preaplin* al cunoașterii, care ne ajută să ne depășim condiția, vorbim de un *exces* al cunoașterii, prin care cei puțini îi controlează pe cei mulți.

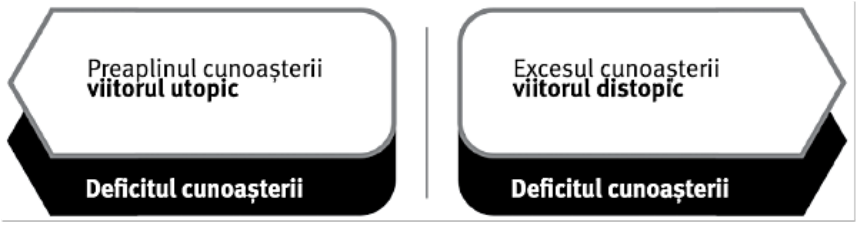


Fig. 3 Inteligența artificială în societatea informațională – un preaplin, un exces sau un deficit al cunoașterii? Ilustrație realizată de autor.

Mai mult decât atât, dincolo de această ambivalență în citirea prezentului și viitorului nostru digital, se conturează și o a treia interpretare, ce pune accentul pe *deficitul* de cunoaștere al deciziilor algoritmice. Cei care propun această versiune examinează numeroasele promisiuni neîndeplinite, erori, limitări, aproximări și mișcări retorice din spatele portretizării inteligenței artificiale ca fiind omniscientă. Câte eșecuri neraportate se ascund în spatele poveștilor de succes ale inteligenței artificiale, despre care citim în mass-media? Cât este forță și cât este coincidență în anecdotele privind cunoașterea superumană pe care algoritmii și Big Data o produc despre noi? Cât este știință, cât este dorință și cât este marketing în promisiunile științei datelor? Avem de-a face, din această perspectivă, nu cu eliminarea sau reducerea erorilor umane, ci cu înlocuirea lor, cu alte tipuri de necunoaștere și de neînțelegere.

Până în secolul XXI, deciziile oamenilor, fie ele corecte sau greșite, bine sau rău intenționate, au fost comunicate și justificate prin vorbe și texte, adică prin media inteligibile uman. Legile, verdictele, opiniile și propunerile de schimbare socială din ultimele secole, chiar și cele mai ferite sau mai malefice, au fost toate exprimate într-o formă pe înțelesul oamenilor – cel puțin pentru observatorii care aveau acces la ele. Lizibilitatea și transparența deciziilor încep, treptat, să se erodeze, pe măsură ce *textul* este înlocuit de combinații cu tot mai multe straturi interdependente de *cod și date*.

extract_num-
ber_and_incr (destination, source) int
*destination; unsigned char **source; { extract_num-
ber (destination, *source); *source += 2; } #ifndef EXTRACT_MAC-
ROS #undef EXTRACT_NUMBER_AND_INCR #define EXTRACT_NUM-
BER_AND_INCR(dest, src) \ extract_number_and_incr (&dest, &src) #endif /*
not EXTRACT_MACROS */ #endif /* DEBUG */ /* If DEBUG is defined, Regex prints
many voluminous messages about what it is doing (if the variable `debug' is nonzero). If
linked with the main program in `iregex.c', you can enter patterns and strings interactively.
And if linked with the main program in `main.c' and the other test files, you can run the al-
ready-written tests. */ #ifdef DEBUG /* We use standard I/O for debugging. */ #include <stdio.h>
/* It is useful to test things that ``must'' be true when debugging. */ #include <assert.h> static int
debug = 0; #define DEBUG_STATEMENT(e) e #define DEBUG_PRINT1(x) if (debug) printf (x) #define
DEBUG_PRINT2(x1, x2) if (debug) printf (x1, x2) #define DEBUG_PRINT3(x1, x2, x3) if (debug) printf
(x1, x2, x3) #define DEBUG_PRINT4(x1, x2, x3, x4) if (debug) printf (x1, x2, x3, x4) #define DE-
BUG_PRINT_COMPILED_PATTERN(p, s, e)\ if (debug) print_partial_compiled_pattern (s, e) #define DE-
BUG_PRINT_DOUBLE_STRING(w, s1, sz1, s2, sz2) \ if (debug) print_double_string (w, s1, sz1, s2, sz2)
extern void printchar(); /* Print the fastmap in human-readable form. */ void print_fastmap (fastmap)
char *fastmap; { unsigned was_a_range = 0; unsigned i = 0; while (i < (1 << BYTEWIDTH)) { if (fastmap[i++])
{ was_a_range = 0; printchar (i - 1); while (i < (1 << BYTEWIDTH) && fastmap[i]) { was_a_range = 1; i++; } if
(was_a_range) { printf ("-"); printchar (i - 1); } } } putchar ('\n'); } /* Print a compiled pattern string in hu-
man-readable form, starting at the START pointer into it and ending just before the pointer END. */ void
print_partial_compiled_pattern (start, end) unsigned char *start; unsigned char *end; { int mcnt, mcnt2; un-
signed char *p = start; unsigned char *pend = end; if (start == NULL) { printf ("(null)\n"); return; } /* Loop over
pattern commands. */ while (p < pend) { switch ((re_opcode_t) *p++) { case no_op: printf ("/no_op");
break; case exactn: mcnt = *p++; printf ("/exactn/%d", mcnt); do { putchar ('/'); printchar (*p++); }
while (--mcnt); break; case start_memory: mcnt = *p++; printf ("/start_memory/%d/%d", mcnt,
*p++); break; case stop_memory: mcnt = *p++; printf ("/stop_memory/%d/%d", mcnt, *p++);
break; case duplicate: printf ("/duplicate/%d", *p++); break; case anychar: printf ("/anychar");
break; case charset: case charset_not: { register int c; printf ("/charset%s", (re_opcode_t) *(p -
1) == charset_not ? "_not" : ""); assert (p + *p < pend); for (c = 0; c < *p; c++) { unsigned bit;
unsigned char map_byte = p[1 + c]; putchar ('/'); for (bit = 0; bit < BYTEWIDTH; bit++) if
(map_byte & (1 << bit)) printchar (c * BYTEWIDTH + bit); } p += 1 + *p; break; } case beg-
line: printf ("/begline"); break; case endline: printf ("/endline"); break; case on_failure_-
jump: extract_number_and_incr (&mcnt, &p); printf ("/on_failure_jump/0/%d", mcnt);
break; case on_failure_keep_string_jump: extract_number_and_incr (&mcnt, &p); printf
("/on_failure_keep_string_jump/0/%d", mcnt); break; case dummy_failure_jump: ex-
tract_number_and_incr (&mcnt, &p); printf ("/dummy_failure_jump/0/%d", mcnt); break;
case push_dummy_failure: printf ("/push_dummy_failure"); break; case may-
be_pop_jump: extract_number_and_incr (&mcnt, &p); printf
("/maybe_pop_jump/0/%d", mcnt); break; case pop_failure_-
jump: extract_number_and_incr (&mcnt, &p); printf ("/pop_-
failure_jump/0/%d", mcnt); break; case jump_past_alt:
extract_number_and_incr (&mcnt, &p); printf ("/-

Fig. 4 Conflictul dintre inteligibilitatea umană și complexitatea sistemelor de cod și date se amplifică. Ilustrație de Gordon Johnson, publicată pe Pixabay.

Sistemele algoritmice bazate pe învățare automată și date vaste sunt deja prea complexe pentru a fi citite de un om sau de un colectiv de oameni. Deciziile lor pot fi inspectate sau testate, dar nu neapărat înțelese. Inteligența artificială explicabilă (*explainable AI*) a devenit deja un domeniu de studiu substanțial. Acesta este dedicat tocmai inteligibilității deciziilor inteligențelor artificiale și provocării de a crea modele IA *explicabile*.

Deși calculatoarele inspiră certitudine și predictibilitate, sistemele actuale de învățare automată care procesează date vaste sunt într-atât de complexe și de schimbătoare, încât produc rezultate surprinzătoare, neinteligibile și dificil de retrasat într-un parcurs retroactiv. Big Data includ adesea erori factuale, care sunt însă costisitor de detectat și de remediat la nivelul unui individ anume. Prin acumulare și compunere de erori și aproximări succesive, mascate de straturi de complexitate și de secretul drepturilor de proprietate intelectuală, inteligențele artificiale pot greși. Ele se află însă la adăpost de a fi detectate și, mai mult decât atât, nu pot fi trase la

răspundere. Eficacitatea presupusă a deciziilor algoritmice în evitarea greșelilor omenești este dublată, așadar, de o umbră a erorilor și arbitrariului non-uman. Rămâne de văzut în ce măsură și prin ce metode idealul transparenței decizionale mai poate fi urmărit într-o societate bazată pe justificări ascunse în măruntaiele centrelor de date și supercomputerelor care rulează modelele ce ne clasifică și ne prezic.

Pe măsură ce oamenii își pierd pozițiile privilegiate din lanțurile decizionale care ne sfătuiesc, ne recrutează, ne etichetează și ne sancționează, apar situații de derapaj. Cei care ajung să sufere costuri de neimaginat sunt, cu precădere, indivizii și diversele comunități umane. Desigur, discriminarea, precaritatea sau excluziunea socială nu au fost inventate în epoca algoritmilor. La originea lor se află conflicte și instituții sociale generatoare de inegalitate, recreate, atât prin reproducere, cât și prin schimbare, timp de secole. În ce măsură traducerea atributelor individuale și a deciziilor umane în variabile prelucrate digital va consolida sau va corecta aceste procese istorice de vulnerabilizare? Ne aflăm în situația remedierii unor părtiniri trecute, dar și a apariției unor noi forme de abuz. Aceste devieri sunt fără precedent prin anvergura și viteza la care pot fi propagate global și prin dificultatea de inspectare a proceselor ce le produc.

Pe măsură ce avansăm în secolul XXI, ne reîntoarcem spre dilemele evidențiate cu douăzeci și cinci de secole în urmă de Platon, în *Republica*. Este dreptatea doar o deghizare pentru justificarea intereselor celor puternici? Sau dreptatea este un echilibru instabil al intereselor cetățenilor egali în demnitate și drepturi? Cum afectează noua reîmpărțire a cunoașterii și a tehnologiilor de control egalitatea dintre participanții la contractul social?

Discuția despre tehnologii și digitalizare adesea maschează dezbaterea politică și etică privind distribuirea puterii. Tehnologiile nu ne influențează doar direct și nemijlocit, inginerește, ci sub forma unor produse. Aceste produse sunt create și dezvoltate de organizații variate, de la start-upuri și ONG-uri, la țări și mari corporații transnaționale, cu valori de piață comparabile cu PIB-ul unor întregi state. Companiile Big Tech nu mai creează, în principal, produse, ci noi infrastructuri și piețe pe care circulă, lucrează și oferteză milioane de alți creatori ai unor produse mai mici.

Care mai este rolul și puterea individului – în ipostaza de consumator sau de cetățean, de angajat, liber-profesionist sau antreprenor – într-o lume în care opțiunile de vot, de business, de informare sau de distracție sunt formulate și mediate nu doar de organizațiile tipice secolului XX, ci și de marii coloși tehnologici ce-și urmăresc, simultan, propriile interese?

Într-o lume tot mai interdependentă, observăm cum, paradoxal, ideologiile individualismului – de sorginte liberală și mai ales neoliberală – câștigă relevanță. Valorizarea individualismului este amplificată de *personalizarea* conținuturilor digitale, oferită prin marketingul predictiv și algoritmii de recomandări. Aplicațiile, rețelele, comunitățile online ne oferă o multitudine fascinantă de oglinzi digitale în care să ne reflectăm și să experimentăm cu fiecare fațetă a identității noastre. În calitate de utilizatori, prin apelul la avataruri și *personae* digitale, participăm în fluxuri continue de interacțiuni prin care ne testăm și ne rafinăm manifestarea virtuală a sinelui. Totuși, în 2019, mai puțin de două treimi din traficul de pe internet a fost generat de oameni, iar aproximativ 37% a fost generat de boți[3]. Aceștia servesc interesele unor organizații, fie ele comerciale, politice sau statale. Aceleași interese stau și la baza existenței multora dintre platformele și infrastructurile digitale pe care circulăm în lumea virtuală. Observăm, așadar, o redefinire radicală a puterii individuale de acțiune, ba chiar a noțiunii de *actor* în viața publică.

În paginile următoare vom discuta felul în care tehnologiile digitale actuale transformă autonomia și controlul de care dispunem, pentru a ne gestiona viața atât la nivel individual, cât și colectiv. Vom porni de la perspectiva *preaplinului* cunoașterii, examinând poveștile de succes și promisiunile noii ordini digitale. În capitolele următoare, preaplinul va fi discutat ca o formă de *exces*, urmat de reversul medaliei, respectiv *deficitul* cunoașterii algoritmice și implicațiile sale. Vom revedea, așadar, din mai multe unghiuri, speranțele de emancipare individuală și socială prin digitalizare, inteligențe artificiale și Big Data, precum și controversele create. Sper ca lectura cărții să deschidă pauze periodice de reflecție, conturând mai bine universul imaginativ în care ne povestim viitorul, cuprinzând cu ochiul minții colosalele transformări care se petrec aici și acum. Deși știrile privind inteligența artificială și capitalismul datelor

punctează periodic peisajul informativ global și românesc, ele se pierd, adesea, în noianul de noutăți politice, economice, sportive care ne captează atenția zi de zi. Prinși în actualitate, riscăm să nu vedem pădurea sau chiar jungla algoritmică din spatele copacilor. Un viitor imaginat poate fi mai ușor actualizat, în funcție de opțiunile noastre colective. Vom discuta alegerile făcute în trei modele diferite ale guvernării, în Statele Unite ale Americii, Uniunea Europeană și China. Deși poate părea că zarurile sunt aruncate doar de giganții tehnologici, Europa continuă să construiască repere morale esențiale pentru societatea puterii algoritmice, iar viitorul este deschis.

---

1. Mai multe detalii despre program sunt disponibile la www.innovationlabs.ro.

2. Conform datelor privind produsul intern brut estimate de Banca Mondială și de Națiunile Unite, sintetizate pe pagina Wikipedia: https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal).

1. Tehnologii digitale, capitalism și modele sociale: SUA, UE și China

Omul sfințește locul, nu locul pe om.

Tehnologiile digitale, inclusiv inteligența artificială sau învățarea automată, nu sunt intrinsec bune sau rele. Impactul lor social depinde de felul în care sunt implementate în produse și servicii, de strategiile organizațiilor private sau publice care le vor folosi, precum și de reacțiile, uneori neprevăzute, ale utilizatorilor.[1] Modelul social în cadrul căruia o tehnologie ajunge să fie implementată și propagată este, de asemenea, foarte important, oferind și limitând resursele cu care diferiți actori sociali intră în negociere.[2] Puterea relativă pe care o au platformele, clienții sau lucrătorii, anvergura circulației și centralizării datelor personale și riscurile tipice la care sunt expuși indivizii depind, în mare măsură, de normele culturale și legale ale unei societăți date.

În încercarea de a stabili un cadru de comparație, putem porni de la trei modele sociale alternative. Acestea sunt relevante pentru evoluția curentă a tehnologiilor digitale. Am în vedere modelul social american, cel european și cel chinez. Să examinăm aceste modele pe trei dimensiuni: justiție, inegalitate și deliberare socială. Voi discuta apoi implicațiile acestor modele pentru impactul tehnologiilor digitale, cu precădere privind inteligența artificială și utilizarea datelor masive.

Autonomia sferelor justiției

Banii n-au miros.

Michael Waltzer, filozof politic și intelectual public american, profesor la Institutul de Studii Avansate din Princeton, observa că fiecare societate împărtășește valori multiple, care sunt ireductibile una la cealaltă, precum prestigiul, educația, sănătatea, avuția, puterea politică, timpul liber, hrana

sau arta.[3] Nu există o listă standard a obiectelor materiale sau imateriale prețuite de o anumită comunitate politică.

Pentru a afla ce valori sunt importante într-o societate la un moment dat, trebuie să studiem empiric și să interpretăm deciziile și practicile acelei societăți de-a lungul istoriei sale. Pentru fiecare dintre aceste valori, o societate va considera că o anumită distribuție este preferabilă alteia. De exemplu, putem crede că este bine ca puterea politică să fie distribuită inegal între cetățeni, adică unii dintre aceștia să aibă funcții de conducere, iar alții să fie subordonați. De asemenea, putem crede că un anumit nivel al inegalității veniturilor și averilor este o sursă a bunăstării comune.

Totuși, multe dintre societățile contemporane consideră că accesul la justiție ar trebui să fie distribuit egal, independent de avere sau de puterea politică. Totodată, preferințele pentru alocarea accesului la sănătate sau la educație sunt, în ziua de azi, mai egalitare decât preferințele pentru alocarea accesului la bunuri de lux.

Distribuțiile ideale se pot schimba rapid, în funcție de impactul transformărilor sociale. De exemplu, odată cu școala online, introdusă de pandemia Covid-19, a devenit necesar ca fiecare copil să aibă acces la internet și la un dispozitiv digital de tip tabletă sau laptop pentru a putea participa la educație. Anterior acestui eveniment, inegalitățile digitale erau considerate grave, dar parțial justificabile și nu neapărat intolerabile.

Ideal vorbind, observă Waltzer, într-o societate dată, fiecare valoare ar trebui să fie distribuită indivizilor conform logicii proprii. Asta presupune ca nivelul la care sunt situați oamenii pe scala altor valori să nu conteze atunci când se stabilește distribuția unei anumite valori. De exemplu, notele școlare nu ar trebui să fie acordate în funcție de veniturile sau influența politică a familiei elevului. Acesta este idealul etic și politic al egalității complexe, în care diferitele sfere ale dreptății distributive nu se influențează unele pe altele. Transferul ilegitim de acces dintr-o sferă într-alta este considerat corupție sau discriminare.

Astfel, în multe societăți actuale, este considerat incorect ca oamenii cu mai multă avere să fie privilegiați în obținerea diplomelor, în actul de justiție, în prestigiul artistic sau în accesul la transplanturi și alte tratamente medicale de urgență. Totuși, este adesea acceptată o condiționare între

bunăstarea economică și accesul la servicii medicale, educaționale sau culturale mai puțin necesare.

Pentru fiecare societate, o întrebare fundamentală care definește modelul social este: Cât de autonome sunt diversele sfere ale dreptății distributive în raport cu avuția? Mai concret, cât de mult depinde accesul la funcții publice ori la servicii de sănătate sau educaționale de resursele financiare ale unei persoane? Cu alte cuvinte, *în ce măsură capitalismul subsumează logicile de distribuție ale altor valori*? Această întrebare este importantă, deoarece afectează în mod direct impactul social al tehnologiilor.

Toate produsele tehnologice promit să ne conducă spre o lume mai bună, să rezolve nevoile populației prin scăderi ale prețurilor, sporiri ale accesului la resurse și fluidizarea interacțiunilor. Totuși, în practică, oare nu vor servi aceste produse digitale exclusiv, sau dominant, intereselor capitalului ce guvernează firmele care îl produc? Sau vor fi avute în vedere și interesele altor participanți, precum angajații, micii antreprenori sau utilizatorii individuali? Cine va extrage cea mai mare parte din valoarea creată de aceste noi produse și piețe?

Un exemplu al acestei tensiuni constă în controversa legată de rolul avut de Facebook în polarizarea socială, amplificarea știrilor false și a discursurilor încărcate de ură. Scandalul Cambridge Analytica[4] a adus în prim-planul atenției riscul manipulării și polarizării politice prin personalizarea tot mai precisă a mesajelor electorale difuzate prin Facebook. Controversa a luat amploare în 2021, când a devenit tot mai clară tensiunea dintre profitabilitatea enormă a companiei, bazată pe reclamele consumate de utilizatori, și posibilele soluții de reducere a discriminării și dezinformării. În competiția dintre profit și spirit civic, oare cine câștigă cursa pe platforma globală care informează și conectează aproape trei miliarde de utilizatori?[5] Răspunsul, cel puțin deocamdată, pare să fie profitul.

În martie 2021, Karen Hao a realizat un reportaj extensiv pentru *MIT Technology Review*. Ea a explicat felul în care tehnologiile Facebook, cuplate cu strategia de creștere agresivă (vezi figura 5) și cu ale opțiuni ale managementului, conduc la proliferarea în continuare a știrilor false[6]. Hao a urmărit activitatea inginerilor de la Facebook, mai ales a celor implicați în

filtrarea dezinformării și în prevenirea discriminării. Ea a studiat, în special, strategia lui Mark Zuckerberg de a nu aliena publicul de orientare republicană din SUA. Jurnalista a observat că deciziile politice luate de conducerea Facebook au condus la un conflict între cele două obiective de responsabilitate socială asumate de Facebook, anume reducerea discriminării și evitarea dezinformării. Cum s-a petrecut acest lucru? În primul rând, trebuie să înțelegem cum funcționează algoritmii care decid ce mesaje vedem, fiecare dintre noi, în fluxul nostru personalizat de pe Instagram sau Facebook. Aceste mesaje sunt împinse în față, pentru a ne capta atenția, prin modele algoritmice care învață continuu. Sursele de antrenare și perfecționare pentru acești algoritmi au la bază reacțiile noastre și ale celorlalți utilizatori. Ce anume ne face să petrecem mai mult timp pe platformă și să fim mai implicați emoțional? Mai specific: ce feluri de mesaje ne țin lipiți de ecran și ne determină să dăm *share* și *like*? Totodată, ce fel de grupuri sunt atractive pentru noi și ne fac să vrem să ne alăturăm lor?

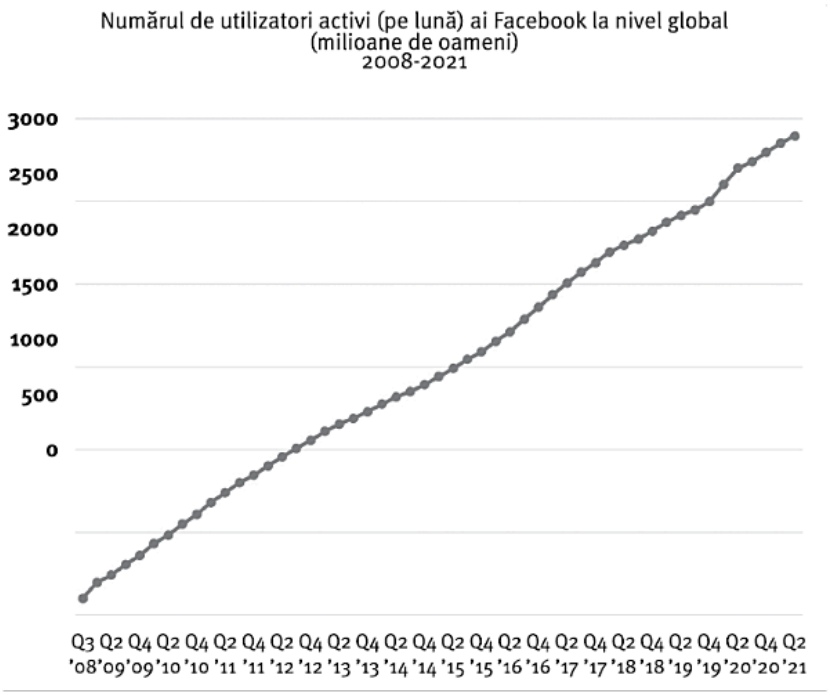


Fig. 5 Creșterea numărului de utilizatori lunar activi ai Facebook, la nivel global, 2008–2021. Grafic realizat de autor pe baza datelor Statista.

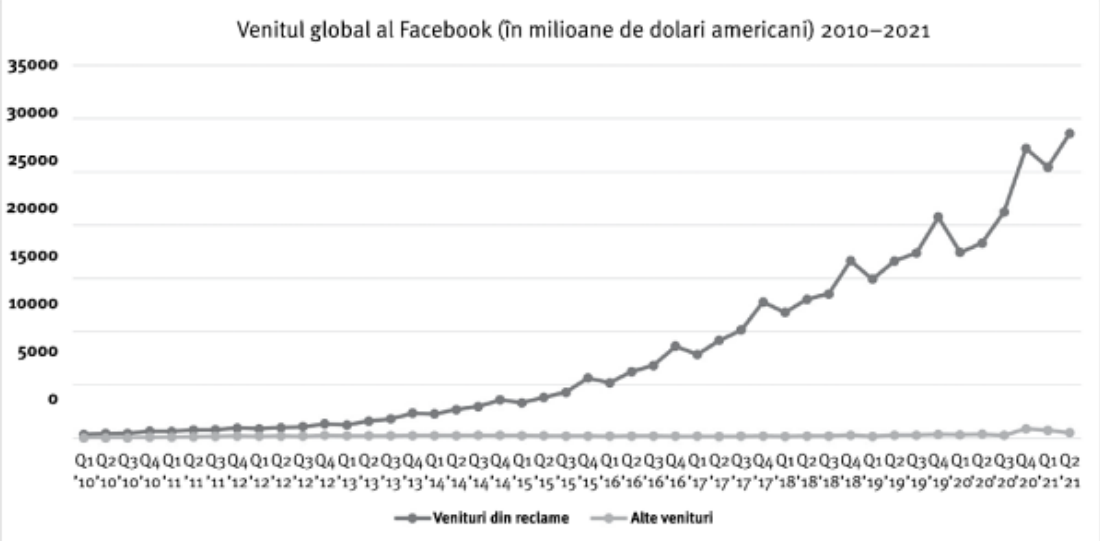


Fig. 6 Creşterea veniturilor din reclame şi alte surse pentru Facebook. Grafic realizat de autor pe baza datelor Statista.

Algoritmii prelucrează continuu urmele digitale pe care noi şi alţi utilizatori similari nouă le lăsăm prin clickurile pe care le dăm pe platformă şi selectează conţinutul cu probabilitatea cea mai mare de a ne captiva. *Modelul de afaceri* al Facebook solicită un timp cât mai mare, tot mai mare, petrecut de tot mai mulţi oameni în aplicaţie. Creşterea este considerată scopul care scuză mijloacele[7] – şi, într-adevăr, Facebook a crescut continuu (vezi figurile 5 şi 6). Cu cât stăm mai mult cu ochii în ecran şi cu inima lipită de *feed*, cu atât suntem mai bine ţintiţi de reclamele care asigură profitul substanţial şi în permanentă ascensiune al companiei[8]. Din păcate, multe dintre ştirile false şi mesajele care incită la indignare şi la ură au un potenţial foarte mare de a atrage publicul. Aceasta se întâmplă deoarece, pe scurt, astfel de ştiri şi de mesaje pot fi mai interesante şi mai emoţionante decât afirmaţii reale şi imparţiale difuzate pe un ton ştiinţific. Prin urmare, există o motivaţie comercială puternică pentru propagarea acestor mesaje care destructurează dialogul rezonabil în societate.

În reportajul său, Karen Hao arată cum modelele făcute pentru a filtra acest conţinut toxic sunt sistematic date la o parte, dacă se dovedeşte că scad atractivitatea platformei pentru utilizatori – ceea ce, adesea, se întâmplă. Cu alte cuvinte, Facebook alege, în cele din urmă, acei algoritmi care vor stimula creşterea numărului de utilizatori, implicarea emoţională a

oamenilor și timpul tot mai mare petrecut pe platformă, în dauna promovării adevărului sau a echilibrului emoțional.

Atunci când Facebook a fost acuzat de președintele Donald Trump că discriminează publicul republican, Mark Zuckerberg a mobilizat echipa de combatere a discriminării pentru a asigura o vizibilitate egală a mesajelor preferate de democrați și de republicani. În practică, această politică de egalizare a vizibilității a condus la proliferarea știrilor false. Motivul pentru care știrile false s-au înmulțit are în vedere faptul că, statistic, ele sunt mai des transmise de utilizatorii cu înclinații republicane – fie că a fost vorba despre atacul asupra Capitoliului, despre negarea gravității pandemiei Covid-19 sau despre promovarea discursurilor antivaccin și antimască.[9]

Aceeași diagnoză a fost repetată de Frances Haugen, din poziția de avertizor de integritate (*whistleblower*), în toamna anului 2021.[10] Haugen a lucrat la Facebook în calitate de manager de produs în echipa de integritate civică, responsabilă tocmai cu remedierea problemelor de dezinformare și polarizare. Ea s-a angajat la Facebook animată de o motivație personală: un prieten foarte bun fusese, treptat, atras în cercuri tot mai conspiraționiste și mai rasiste, iar prietenia lor a fost distrusă de această dinamică.[11] Frances Haugen a fost indignată de ceea ce a considerat că sunt practici imorale ale companiei, care strică tot ceea ce aceasta face bine și ar putea face și mai bine. Ea a dezvăluit numeroase rapoarte interne și alte documente ale companiei către *The Wall Street Journal* și apoi către SEC (Comisia Americană pentru Bursă și Valori mobiliare), fiind ulterior audiată de Senatul SUA. Aceste documente dovedesc, conform lui Haugen, faptul că Facebook pune pe primul plan profitul, în dauna respectării promisiunilor făcute de a ajuta oamenii să se conecteze unii cu alții și să întărească țesătura socială. Întrebarea pe care o ridică atât Karen Hao, cât și Frances Haugen este: Ce limite și ce promisiuni făcute publicului este dispusă o companie să încalce în scopul expansiunii? Mai mult, ce fel de încălcări ale promisiunilor sunt dispuși cetățenii și consumatorii să accepte?

Economistul american Milton Friedman a lansat în imaginația publică o nouă definiție a moralității corporative, definiție care rămâne influentă și în ziua de azi. Conform teoriei acționarului (*shareholder theory*) sau doctrinei Friedman, „responsabilitatea socială a afacerilor constă în sporirea

profitului“[12]. Cu alte cuvinte, interesele forței de muncă și ale celorlalți participanți într-o firmă sunt neglijabile față de interesele capitalului. Cine sunt ceilalți participanți? Aceștia includ deținătorii terenului sau clădirilor, dar și publicul societăților-gazdă, cele care permit funcționarea companiilor, oferă infrastructura și, adesea, asigură piețe de desfacere. În această viziune, firmele sunt vehicule amorale, care au ca misiune doar maximizarea profitului prin orice mijloace tolerate de societatea în care funcționează. Solicitările de împărțire a câștigurilor cu forța de muncă, prin salarizare, și cu societatea-gazdă, prin autoasumarea unor constrângeri morale și prin contribuții la bugetele publice, sunt considerate nelegitime. Mai mult, acestea sunt echivalente unui furt din buzunarul acționarilor. Conform acestei viziuni neoliberale, foarte populară în SUA, acuzațiile lui Haugen par goale de sens: desigur că Facebook încearcă să-și maximizeze clientela și profitul prin orice metode. De ce ne mirăm? Care este problema?

Unul dintre principiile morale esențiale ale societăților democratice este egalitatea în fața legii. Știm că, în practică, oamenii sunt inegali în multe feluri și că persistă privilegii și dezavantaje moștenite din familie, preluate din mediul social în care trăim sau induse de forțe economice care ies de sub controlul nostru. Totuși, credem că legile și regulile trebuie să fie aplicate la fel în cazul tuturor, indiferent de avere, frumusețe, talent sau faimă. Or documentele dezvăluite de Frances Haugen au arătat că exact acest lucru nu se petrece în interiorul Facebook, cu buna știință a managementului. Compania nu aplica aceleași reguli de moderare tuturor. Pentru a nu supăra sau goni utilizatorii cei mai faimoși și mai bănoși, Facebook avea un program de exceptare a acestora de la reguli, numit XCheck.[13] Persoanelor considerate demne de atenția publică, influente sau populare sau cu riscuri de PR le erau permise postări care ar fi fost prompt moderate sau șterse în cazul utilizatorilor obișnuiți. În 2020, lista celor speciali includea nu mai puțin de 5,8 milioane de nume. Folosindu-se de aceste privilegii, persoanele etichetate ca aflându-se deasupra regulilor au postat, printre altele, fotografii nud fără consimțământul femeii pozate, afirmații false legate de tratamentul cancerului, teorii ale conspirației privind scandalul Pizzagate[14] sau vaccinarea și alte mesaje pe care înșiși verificatorii Facebook le etichetaseră ca știri false. Vedem aici cum dorința

de expansiune și profit, prin monetizarea faimei și influenței acestor VIP-uri, contrazice direct un principiu etic esențial al societăților contemporane, împărtășit de cea mai mare parte a indivizilor.

Societățile europene sunt mai puțin dominate de teoria acționarului formulată de Friedman, fiind mai degrabă adeptele teoriei părții interesate (*stakeholder theory*). În această accepțiune, câștigurile produse de o companie trebuie împărțite între toți participanții la procesul care a produs valoarea – inclusiv angajații și societatea-gazdă, nu doar deținătorii de capital. O firmă care încalcă vizibil interesele cetățenilor societății-gazdă, prin amplificarea violenței politice, prin exploatarea emoțională a clienților sau exploatarea financiară a angajaților, este considerată imorală. Dar în SUA, într-o societate în care publicul este obișnuit cu ideea că este firesc și normal ca firmele să stoarcă bani și din piatră seacă, este nevoie de acuzații extrem de grave pentru a ridica semne de întrebare privind deciziile unei companii. Ambele autoare au acuzat managementul Facebook că este pe deplin conștient, din propriile investigații și rapoarte, de daunele emoționale și sociale pe care le produce strategia de creștere a companiei. Cu alte cuvinte, vătămările produse erau factorizate doar ca un posibil cost economic, justificat prin profitul final. Aceste daune calculate se referă chiar și la copii. Studiile de impact au arătat managementului Facebook că timpul petrecut pe Instagram amplifică gândurile depresive și chiar suicidale ale tinerelor adolescente.[15] De asemenea, era cunoscut faptul că țintirea superprecisă a oamenilor cu mesaje stimulante emoțional îi face să fie mai furioși, mai legați de discursuri și grupuri incitatoare la ură. Aceasta se întâmpla în detrimentul încurajării persoanelor de a se conecta cu cei apropiați și dragi lor – care este, de fapt, obiectivul declarat al companiei.[16]

În repetate rânduri, Haugen a văzut cum posibilele soluții cu eficacitate dovedită pentru remedierea acestei alienări au fost date la o parte pentru a nu scădea implicarea utilizatorilor și a nu afecta creșterea profitului. Atacul asupra Capitoliului a fost, pentru ea, picătura care a umplut paharul, mai ales că acesta a survenit tocmai după ce Facebook renunțase la algoritmii care filtrau mesajele politice incitatoare la ură și violență. Cu alte cuvinte, riscurile de lovitură de stat fuseseră anticipate, iar amplificarea violenței împotriva legiuitorilor a fost doar materializarea unei probabilități deja

asumate de Facebook.

Până în 2021, cea mai mare sancțiune primită de Facebook a constat în amenda de cinci miliarde de dolari administrată de Federal Trade Commission (FTC) în iulie 2019. Această amendă vine ca urmare a scandalului Cambridge Analytica, sancționând incapacitatea companiei de a proteja datele personale ale utilizatorilor. Pe de o parte, volumul amenzii a fost neașteptat de mare, substanțial mai ridicat decât orice altă amendă dată până atunci în SUA pe teme de protecție a datelor (vezi figura 7). Dacă ne uităm în Uniunea Europeană, sancțiunea depășește volumul total al amenzilor date până în 2021 pentru încălcările Regulamentului General privind Protecția Datelor GDPR[3], de aproximativ 1,76 miliarde de euro[4] (vezi și cele mai mari amenzi pentru încălcarea GDPR până în 2021, în figura 8). Totuși, această sancțiune a fost, paradoxal, o veste bună pentru companie – ceea ce s-a văzut prin faptul că, după ce a fost anunțată, cotația la bursă a acțiunilor Facebook a crescut cu peste 1%.[17] Partea plină a paharului a constat în faptul ca Facebook nu a trebuit să facă nici o modificare semnificativă în modelul său de business și în tipul de acces la datele utilizatorilor pe care îl permite, contra cost, unor terțe părți.[18] Această amendă a sancționat, de altfel, încălcarea încrederii consumatorilor – mai degrabă decât problemele sociale provocate prin manipularea politică pe care o permite. Cu alte cuvinte, sancțiunea vine într-o logică de afaceri mai mult decât într-o logică a drepturilor omului sau a apărării valorilor democrației. Rămâne de văzut dacă și cum vor fi sancționate marile companii de tehnologie, aflate sub investigație[19] pentru contribuția rețelelor sociale (Facebook, Instagram, YouTube, Twitter și Reddit) la amplificarea violenței premergătoare atacului asupra Capitoliului.[20] Dată fiind puterea financiară enormă a acestor giganți, amenzile, chiar și cele de ordinul miliardelor de euro, riscă să nu conducă la o schimbare a modelului de afaceri. Dimpotrivă, ajung să îl legitimeze ca metodă verificată de maximizare a profitului. Dacă daunele aduse utilizatorilor și societăților pot fi compensate prin amenzi, fără necesitatea schimbării relației cu utilizatorii și cu clienții, aceste plăți vor deveni un rău necesar sau chiar un preț convenabil, achitat pentru continuarea politicilor comerciale incriminate de public, dar atât de profitabile pentru acționari.

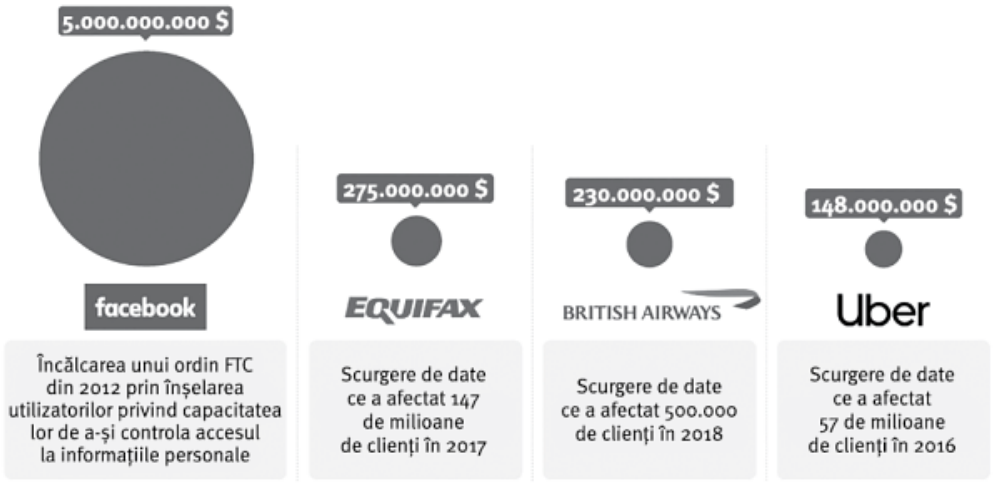

5.000.000.000 $
facebook
Încălcarea unui ordin FTC din 2012 prin înșelarea utilizatorilor privind capacitatea lor de a-și controla accesul la informațiile personale
275.000.000 $
EQUIFAX
Scurgere de date ce a afectat 147 de milioane de clienți în 2017
230.000.000 $
BRITISH AIRWAYS
Scurgere de date ce a afectat 500.000 de clienți în 2018
148.000.000 $
Uber
Scurgere de date ce a afectat 57 de milioane de clienți în 2016

Fig. 7 Amenda administrată de FTC către Facebook în contextul altor sancțiuni pe teme de protecție a datelor. Sursă: Statista.

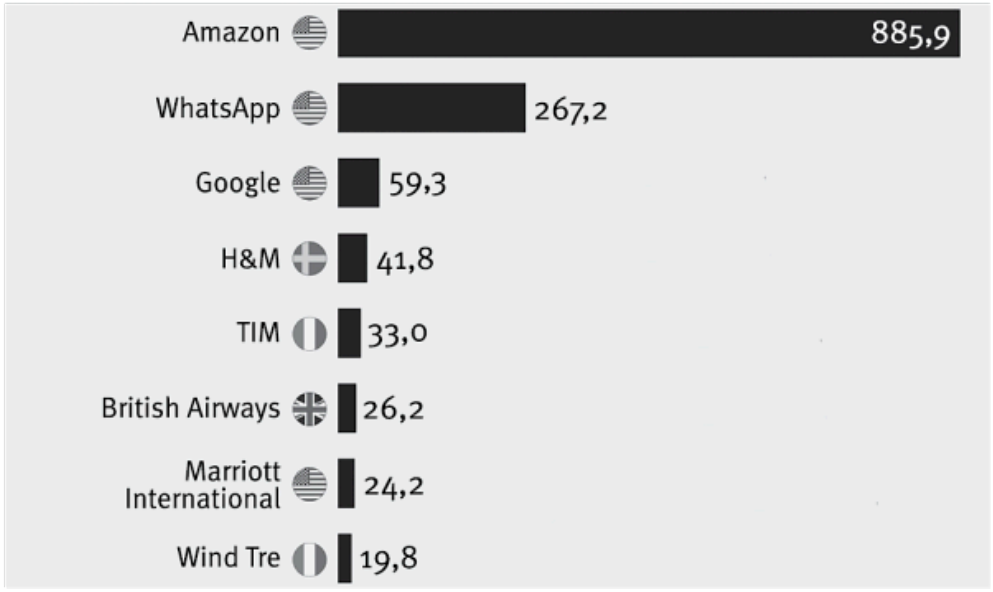


Fig. 8 Cele mai mari amenzi date în Uniunea Europeană pentru încălcarea GDPR (în milioane de dolari). Sursă: Statista

Facebook și Google au modele de afaceri similare, bazate pe personalizarea conținutului și a reclamelor vizionate de utilizatori. De aceea, nu doar Facebook, ci și Google a fost implicată în scandaluri similare legate de instigarea la ură și polarizarea politică, nu atât prin platforma sa de căutare, cât mai ales prin YouTube.

Algoritmul de recomandări al YouTube este responsabil de aproximativ 70% din timpul total de vizionare.[21] Astfel, acest algoritm are o influență

semnificativă atât asupra calității experienței utilizatorilor, cât și asupra profitabilității și impactului social al companiei. Efectele de radicalizare algoritmică ale YouTube au fost documentate și denunțate în repetate rânduri, prin investigații științifice[22] și jurnalistice[23].

La fel ca în cazul Facebook, schimbările aduse de YouTube politicilor de conținut și algoritmilor de recomandare au făcut ca gigantul să își asume, declarativ, obiectivul descurajării radicalizării. Totuși, în practică, a continuat stimularea conținutului incitant[24] care promovează, cu succes, creșterea continuă a numărului utilizatorilor[25] și a profitului din reclame (vezi figura 9).

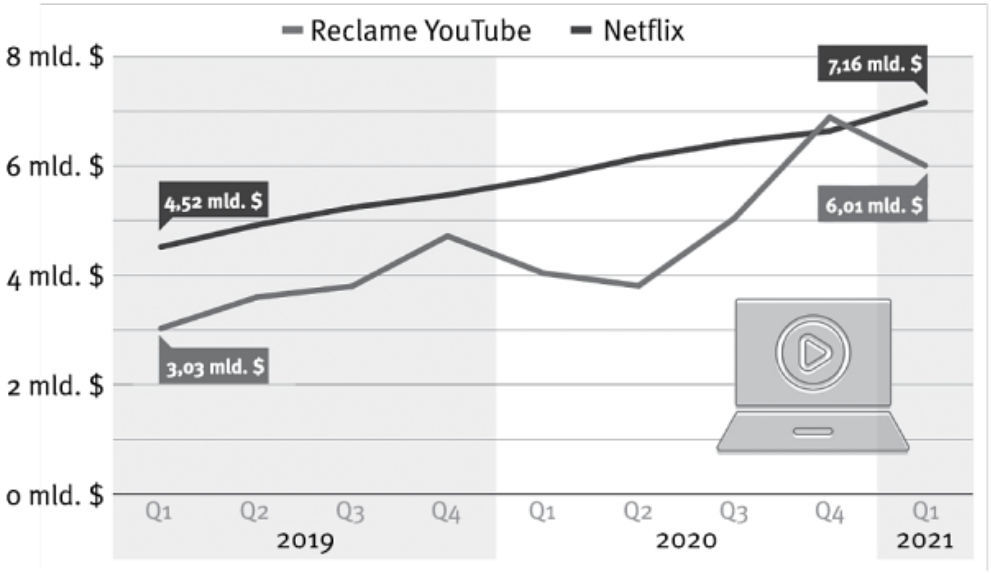


Fig. 9 Creșterea profitului din reclame pentru YouTube a fost accelerată în 2020/2021, atingând nivelul veniturilor Netflix (în miliarde de dolari). Sursa: Statista.

Statele Unite ale Americii oferă, actualmente, un model substanțial diferit de cel al Uniunii Europene privind importanța logicii profitului, iar economia internetului în China este, la rândul său, organizată pe cu totul alte principii.

Începând cu anii ’80, marii deținători de capital din SUA au câștigat din ce în ce mai multă putere. Ei și-au consolidat atât averea, cât și influența asupra politicii și chiar asupra procesului electoral. Pe măsură ce marile corporații au înghițit micii competitori în tot mai multe domenii, logica pieței libere a devenit distorsionată. Marile firme și-au întărit astfel puterea

de a stabili prețuri și de a ridica bariere pentru intrarea competitorilor în piață.

Dacă în SUA interesele capitalului privat – mai ales ale vârfurilor elitei financiare – au ajuns în ultimele decenii să domine vizibil agenda publică și organizarea socială, în China interesul dominant este cel public. Totuși, interesul public este definit doar de Partidul Comunist Chinez – singura organizație politică care guvernează acest stat.

Pentru început, voi ilustra specificul modelului american, comparativ cu cel european și cel chinez, privind dominația capitalismului. Câteva exemple ale diferenței marcante dintre SUA și alte state includ: îndatorarea studenților, costurile pentru serviciile de sănătate, regimul încarcerării și, nu în ultimul rând, finanțarea campaniilor electorale.

Schimbarea structurii pieței forței de muncă în Statele Unite ale Americii, pe parcursul secolului XX, a însemnat că, de la un deceniu la altul, slujbele bine plătite presupuneau tot mai frecvent studii superioare. Totuși, tipul de capitalism practicat în SUA, mai ales începând cu administrația Reagan, a condus la o creștere rapidă a costurilor educației și la scăderea concomitentă a sprijinului financiar oferit de stat.[26] Această schimbare s-a petrecut pe fondul sporirii frecventării instituțiilor educaționale operând pentru profit, în contextul alocării unor resurse tot mai mari din bugetele universităților pentru salariile managementului[27]. Unul dintre efecte a constat în îndatorarea masivă și în continuă creștere a studenților (vezi figura 10).

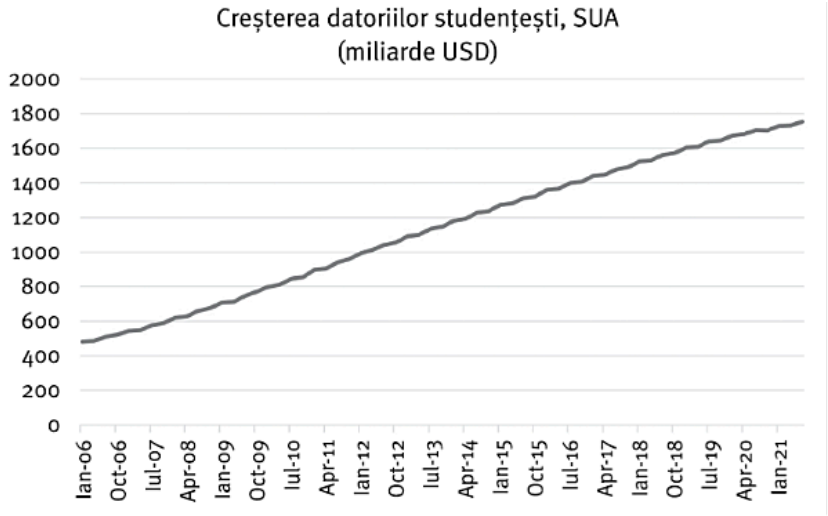


Fig. 10 Creșterea volumului datoriilor acumulate pentru plata studiilor superioare în SUA, 2006–

2021. Grafic realizat de autor pe baza datelor Board of Governors of the Federal Reserve System (US), conform platformei FRED.

Această intensificare a îndatorării a accentuat o criză care afectează întreaga populație, împiedicând atât îmbunătățirea nivelului de trai, cât și transmiterea bunăstării între generații.[28] Dacă în anul 1980 datoriile studențești erau de patru milioane de dolari[29], în trimestrul al doilea al anului 2021 datoria totală acumulată pentru plata studiilor superioare era de aproximativ 1.730 de miliarde de dolari.[30] Estimările arată că aceasta ar putea crește la aproximativ 3.000 de miliarde de dolari până la finalul anului 2030.[31]

Criza îndatorării studențești, specifică societății americane, nu are corespondent în Uniunea Europeană. În UE, o proporție considerabilă a costurilor studiilor universitare este susținută de fiecare stat, nivelul acestor costuri fiind, de asemenea, mai scăzut decât în SUA.

În 2021, China a reacționat preventiv și dur la tendința de creștere a costurilor educației.[32] Începând cu iulie 2021, guvernul chinez a interzis așa-numitele meditații sau lecții private, reducând brusc la zero modelele de afaceri ale platformelor *ed-tech* care găzduiau astfel de cursuri. Legislația chineză urmărește reducerea presiunii financiare pe care părinții o resimt în competiția acerbă a pregătirii copiilor pentru examenele de admitere la liceu și la facultate, precum și reabilitarea, în ochii publicului, a școlilor vocaționale. Vedem cum, în fața unei tendințe vizibile de împovărare a familiilor cu costuri școlare, statul chinez a optat pentru restricții dure. Scopul este acela de a egaliza șansele de succes ale tinerilor prin limitarea relevanței averii părinților și a logicii capitaliste în accesul la oportunități de studiu, examenele de promovare între etape educaționale fiind foarte competitive.

Educația nu este singurul domeniu din SUA în care dominația intereselor capitalului a condus la extragerea masivă de valoare din societate, fiind afectate segmente largi de populație. Un proces similar poate fi observat și în domeniul serviciilor de sănătate. Datorită predominanței motivului profitului față de motivul accesibilității publice, în organizarea serviciilor de sănătate, SUA este o societate unic poziționată în relația dintre costurile sănătății și speranța medie de viață. Practic, finanțarea serviciilor medicale

este cea mai ridicată din lume, dar speranţa medie de viaţă obţinută este semnificativ sub nivelul multor societăţi europene şi de pe alte continente, cu cheltuieli de sănătate mult mai reduse. Momentul în care evoluţia SUA se detaşează de restul ţărilor, amplificând costurile fără o creştere corespunzătoare a speranţei de viaţă, este anul 1983, la începutul administraţiei Reagan, aşa cum poate fi văzut în figura 11.

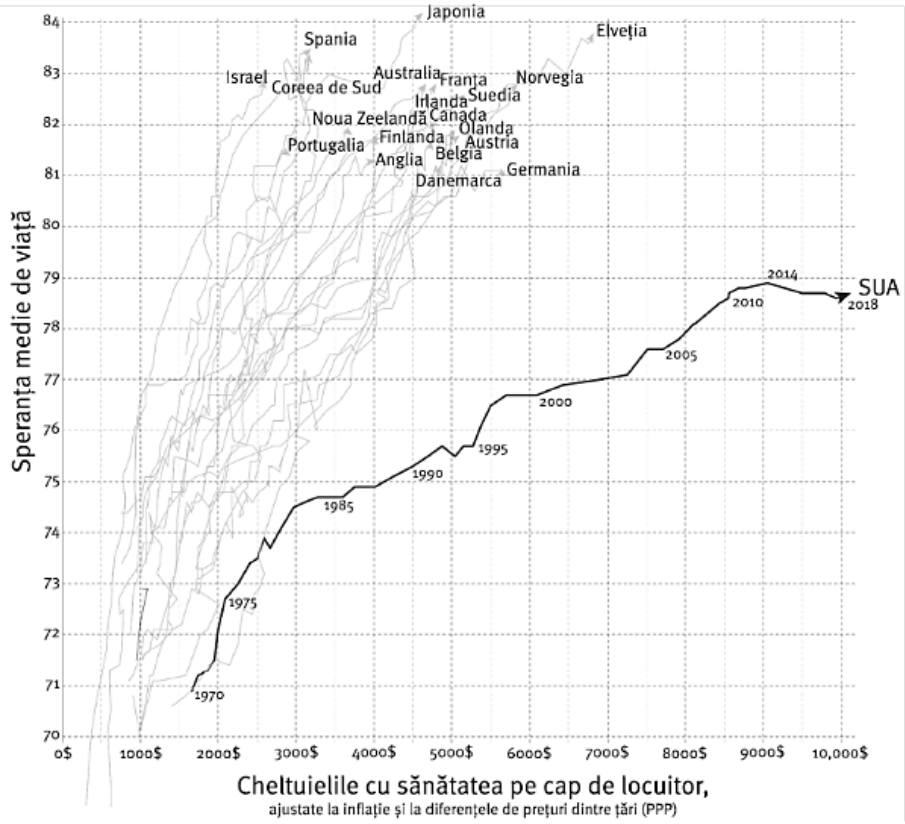


Fig. 11 Relaţia dintre speranţa de viaţă la naştere şi finanţarea serviciilor de sănătate per capita, în evoluţie istorică (1970–2018), comparând SUA cu alte ţări de pe glob. Sursa: Our World in Data.[5]

Criza americană a opioidelor, declanşată prin abuzul analgezicelor bazate pe aceste substanţe, este un alt exemplu în care exacerbarea logicii profitului faţă de celelalte consideraţii şi mecanisme de control a condus la o situaţie fără precedent. După cum putem observa în figura 12, începând cu anii ’90 rata deceselor provocate de supradoze de substanţe opioide a crescut în SUA tot mai rapid. Acest dezastru a condus la aproximativ 500.000 de decese până în toamna anului 2021.[33] La originea sa stau campaniile de marketing agresiv şi manipulativ realizate de companii

farmaceutice[34], în conflict direct cu logica protejării sănătății și interesului public.

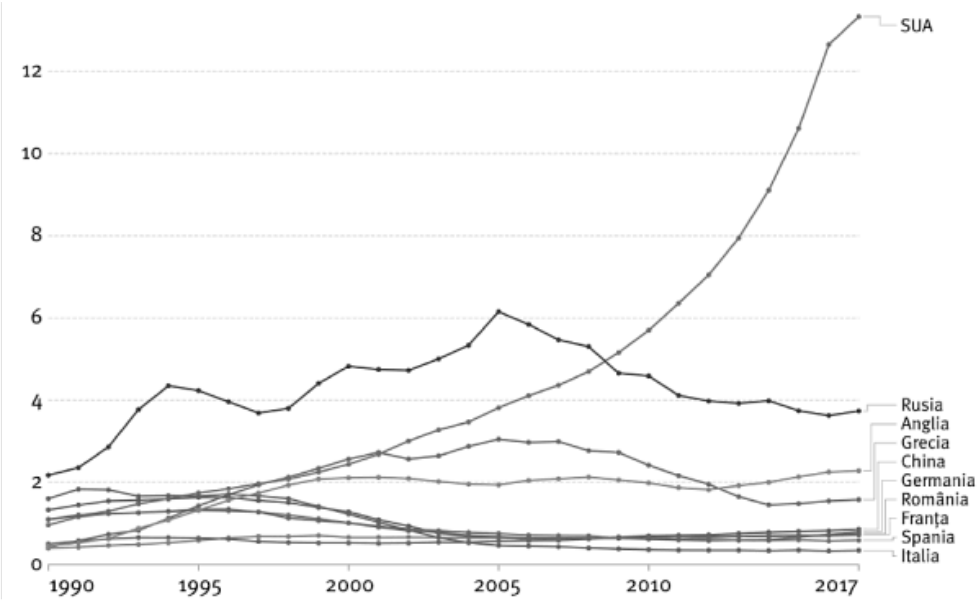


Fig. 12 Rata mortalității în urma abuzului de substanțe opioide în SUA comparativ cu alte state de pe glob. Sursa: Our World in Data.

Exercitarea justiției este un alt exemplu în care dominația logicii capitaliste asupra celorlalte sisteme sociale din SUA este clar vizibilă. Așa-numitul complex penitenciar-industrial (*prison-industrial complex*)[35] se referă la combinația de interese care a condus la privatizarea treptată a multe servicii necesare sistemului penitenciar public, precum servicii legate de hrană, de sănătate, de comunicații, de transport sau de pază, dar și la apariția sistemului de închisori pentru profit.[36] Acest complex a consolidat și continuă să stimuleze, prin motivații comerciale, regimul legal de încarcerare în masă, mai ales a populației de culoare, declanșat prin Războiul împotriva Drogurilor, inițiat de președintele Richard Nixon. Observăm în nivelul extraordinar de ridicat al încarcerării populației americane, comparativ cu celelalte state de pe glob, inclusiv țările europene sau China. De altfel, cele două puncte de inflexiune în creșterea recentă a populației captive în închisori sunt reprezentate de această politică represivă privind drogurile și de reconfigurarea sistemului capitalist sub președintele Ronald Reagan (figura 13).

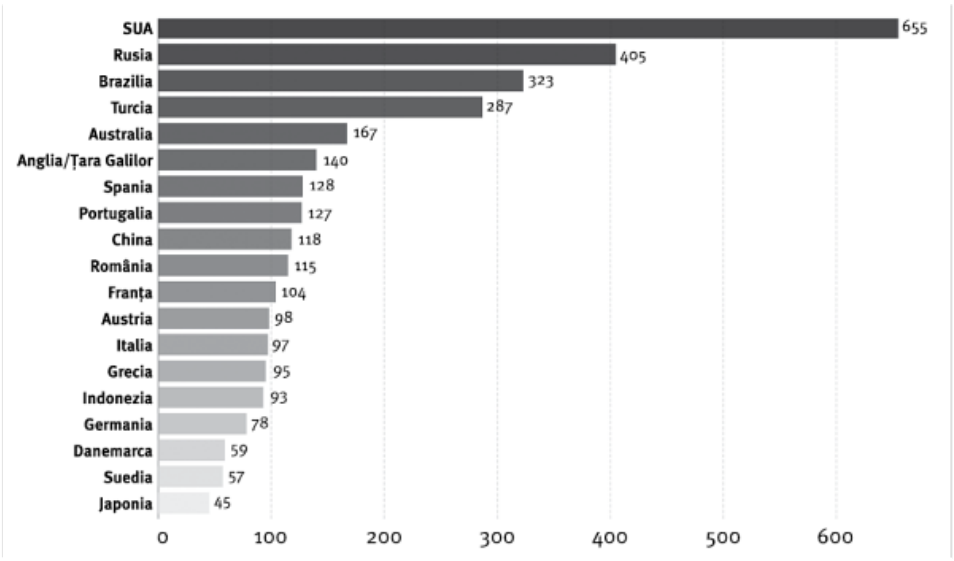


Fig. 13 Rata încarcerării populației în mai multe țări de pe glob (numărul de deținuți raportat la 100.000 de locuitori). Sursa: Our World in Data.

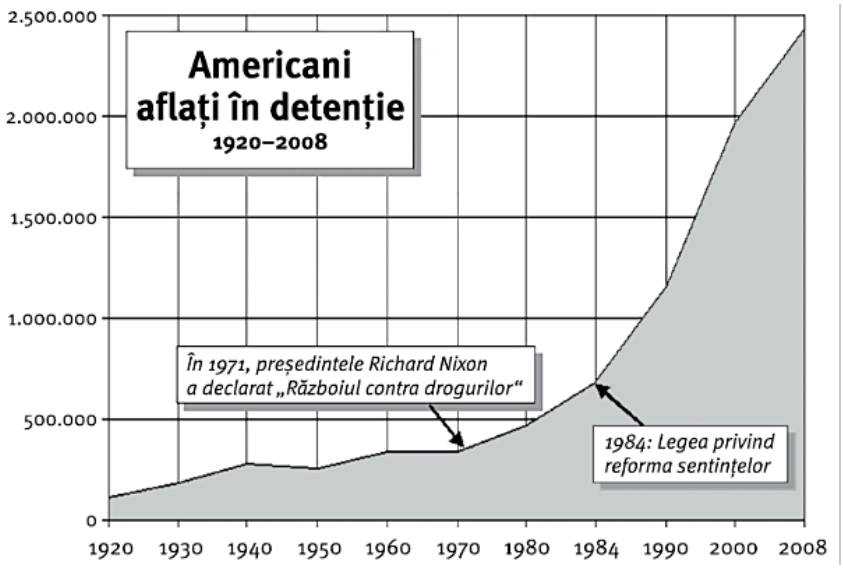


Fig. 14 Evoluția în timp a numărului persoanelor încarcerate în Statele Unite. În 1971, președintele Richard Nixon declară Războiul împotriva drogurilor. În 1984, în timpul președintelui Ronald Reagan, este promulgată o nouă legislație, mai strictă, privind acordarea pedepselor. Sursa: Wikimedia.org.

Nu în ultimul rând, logica preeminenței intereselor capitalului asupra oricăror alte interese și valori sociale, consolidată odată cu anii ’80, a condus și la interferențe între valorile economice și cele ale politicii democratice. Din nou, SUA este un stat unic, printre țările cu nivel de trai ridicat, prin sumele extravagante cheltuite în campaniile electorale.

Decizia Curții Supreme de Justiție din 2010 în cazul Citizens United v. Federal Election Commission a accentuat această interferență. Curtea a echivalat finanțarea campaniilor electorale cu exercitarea dreptului la liberă exprimare. De asemenea, aceeași Curte a declarat că dreptul la liberă exprimare se aplică și organizațiilor, nu doar indivizilor, adică inclusiv corporațiilor, fundațiilor sau uniunilor sindicale. Astfel, aceste organizații au devenit actori politici importanți prin sumele semnificative donate în campaniile electorale. Prin urmare, volumul absolut al cheltuielilor de campanie a crescut (vezi figura 15). În același timp, s-a modificat și structura contribuțiilor[37], ceea ce a dus la sporirea rapidă a finanțărilor independente, externe partidelor (evoluând de la 1% în 1998 la 19% în 2018). Totodată, a crescut proporția sumelor primite de la top 1% al donatorilor externi. De asemenea, a crescut proporția donațiilor mari, în defavoarea micilor donatori.[38]

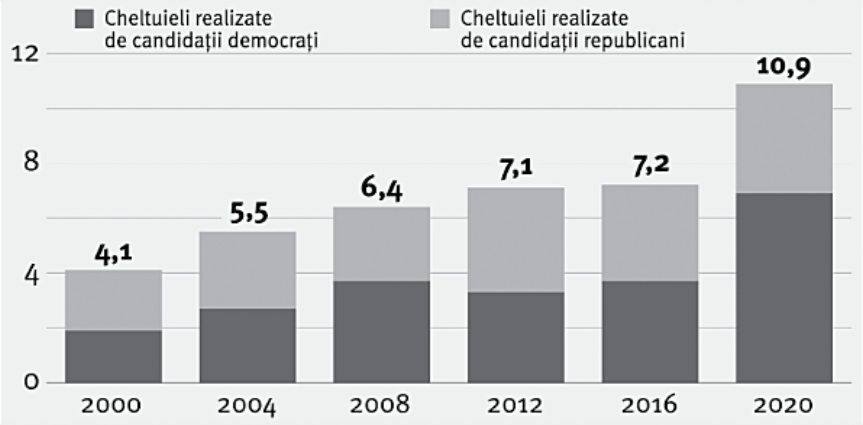


Fig. 15 Evoluția cheltuielilor electorale pentru alegerile parlamentare și prezidențiale din SUA, 2000–2020, incluzând toate sursele de cheltuieli (în miliarde de dolari). Sursa: Statista.

Contrastul cu modelul european de separare a sferelor valorilor și de menținere a autonomiei logicii democratice este vizibil atât prin nivelul absolut al cheltuielilor (vezi figura 16), cât și prin evenimentele majore care au marcat viața politică europeană.

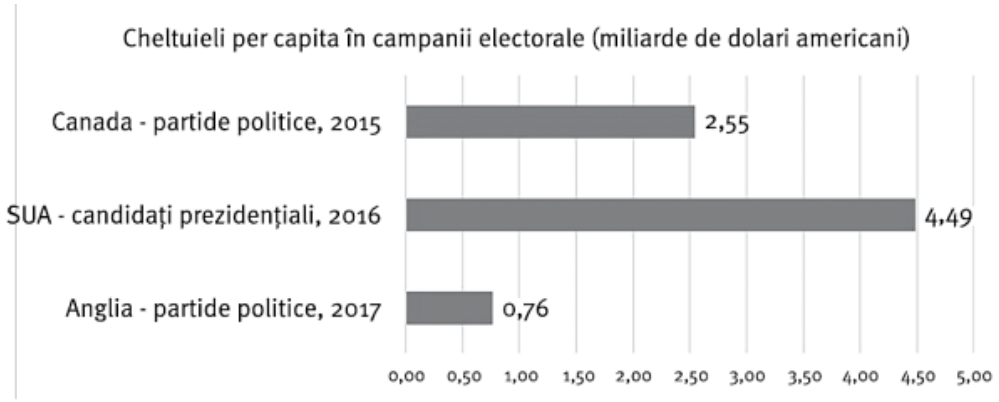

Cheltuieli per capita în campanii electorale (miliarde de dolari americani)
Canada - partide politice, 2015
2,55
SUA - candidați prezidențiali, 2016
4,49
Anglia - partide politice, 2017
0,76
0,00 0,50 1,00 1,50 2,00 2,50 3,00 3,50 4,00 4,50 5,00

Fig. 16 O comparație a cheltuielilor electorale per capita din SUA, Regatul Unit și Canada. Ggrafic realizat de autor pe baza datelor publicate de Horncastle (2020)[39].

Bunăoară, în septembrie 2021, Nicolas Sarkozy, președintele Franței în perioada 2007–2012, a fost condamnat la un an de închisoare pentru depășirea limitei de finanțare a campaniei electorale. În 2012, plafonul întregii campanii prezidențiale era de aproximativ 26,4 milioane de dolari, pentru ambele tururi de scrutin. Președintele Sarkozy – după cum s-a dovedit ulterior – a cheltuit aproximativ 50 de milioane de dolari.[40] Acest exces însă nu a rămas nici neobservat, nici nepedepsit.

Un scandal masiv privind finanțarea campaniilor electorale a fost responsabil și de declinul guvernării cancelarului Helmut Kohl.[41] Realizările sale politice sunt larg recunoscute, precum reunificarea Germaniei, susținerea integrării europene și soluționarea Războiului din Bosnia. Totuși, finanțarea ilegală a campaniilor sale nu a fost trecută cu vederea.

Prin sancțiunile impuse, observăm că eforturile statelor europene de control al plafonului maxim de cheltuieli electorale constituie un pilon important pentru menținerea autonomiei procesului democratic. Această cenzură se remarcă cu atât mai mult cu cât ea acționează împotriva interferențelor logicii de privilegiere a intereselor capitalului.

Să examinăm, în continuare, o a doua dimensiune de diferențiere a modelului american față de cel european și de cel chinez, și anume toleranța față de inegalitatea socială. Aceste comparații sunt interesante mai ales prin întrebările privind tehnologia digitală, formulate în umbra graficelor: în ce măsură aceste modele sociale influențează impactul social al produselor

digitale? Sunt oare tehnologiile nişte forţe exportatoare ale principiilor de organizare socială din ţările în care sunt generate? Pot statele exercita o influenţă semnificativă asupra tehnologiilor disruptive care intervin în organizarea vieţii lor sociale?

Regimuri ale inegalităţii

Omul face haina, şi nu haina pe om.

Thomas Piketty, economist francez, profesor la École des Hautes Études en Sciences Sociales din Paris şi la London School of Economics, discută în cartea Capital et idéologie diferitele modele sociale de generare, reglementare şi justificare a inegalităţilor. El numeşte aceste modele „regimuri ale inegalităţii. Piketty observă cum, începând cu anii 1980, inegalităţile au crescut în toate regiunile de pe glob, cu excepţia celor în care se aflau deja la un nivel ridicat.[42] Creşterea a survenit pe fondul unor eşecuri ideologice ale stângii şi ale succeselor ideologice ale dreptei de orientare neoliberală (vezi figura 17). Schimbarea a fost semnificativ mai pronunţată în SUA faţă de Europa.[43] Raportându-ne la inegalităţile de ordin economico-financiar, putem observa aceste transformări şi în figura 17, unde este vizibil punctul de inflexiune din anii ’80. Remarcăm şi specificul Statelor Unite, clar detaşate de modelul european, acesta din urmă fiind caracterizat, în linii mari, printr-un nivel mai redus al inegalităţilor.

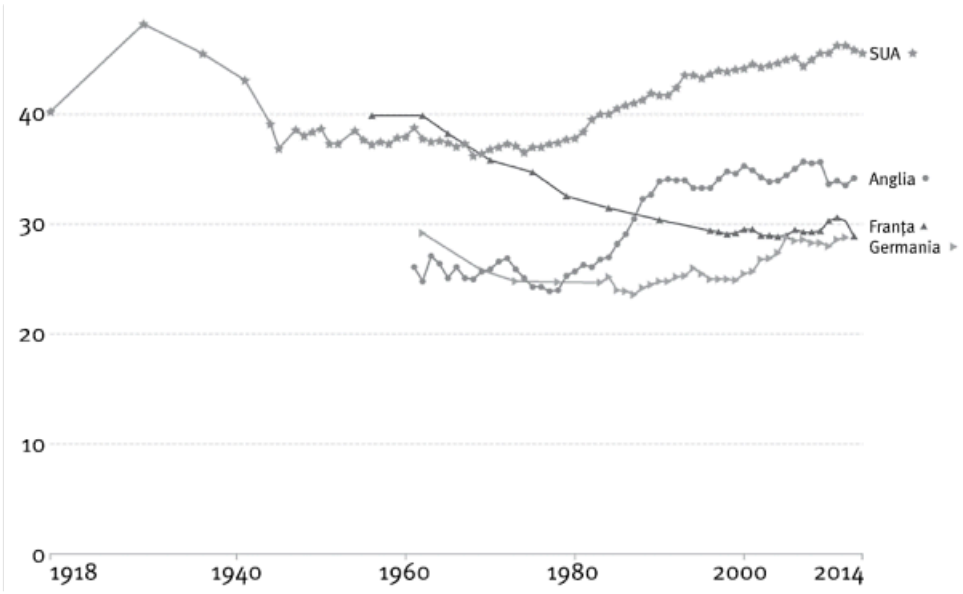


Fig. 17 Evoluția inegalității veniturilor pentru SUA, comparativ cu Anglia, Franța și Germania, 1918–2014. Sursa: Our World in Data.

Această perioadă a anilor 1980–1990 coincide cu anii apariției internetului și declanșării transformării sociale prin digitalizare. Prin urmare, la nivel global și, mai ales în economia și societatea din SUA, trei logici au început să se amplifice reciproc: (1) legitimarea inegalității printr-o ideologie a individualismului suprem și a statului minimalist, (2) accelerarea schimbării tehnologice și economice prin internet și digitalizare și (3) circulația globală a noilor idei și modele de business.

În figura 18, o examinare a evoluției concentrării venitului în mâinile celor 1% cei mai bogați oameni din societate arată o tendință similară. Interesant este că, în ciuda aversiunii ideologiei comuniste față de inegalitate, China nu face excepție de la tiparul de creștere rapidă a inegalităților. Pornind de la un nivel mai scăzut decât statele europene, concentrarea venitului din China în acel mic procent crește mai rapid decât în Vest, ajungând în 2015 la un nivel superior Franței și Germaniei (dar similar celui din Regatul Unit). Statele Unite își păstrează și aici profilul distinctiv.

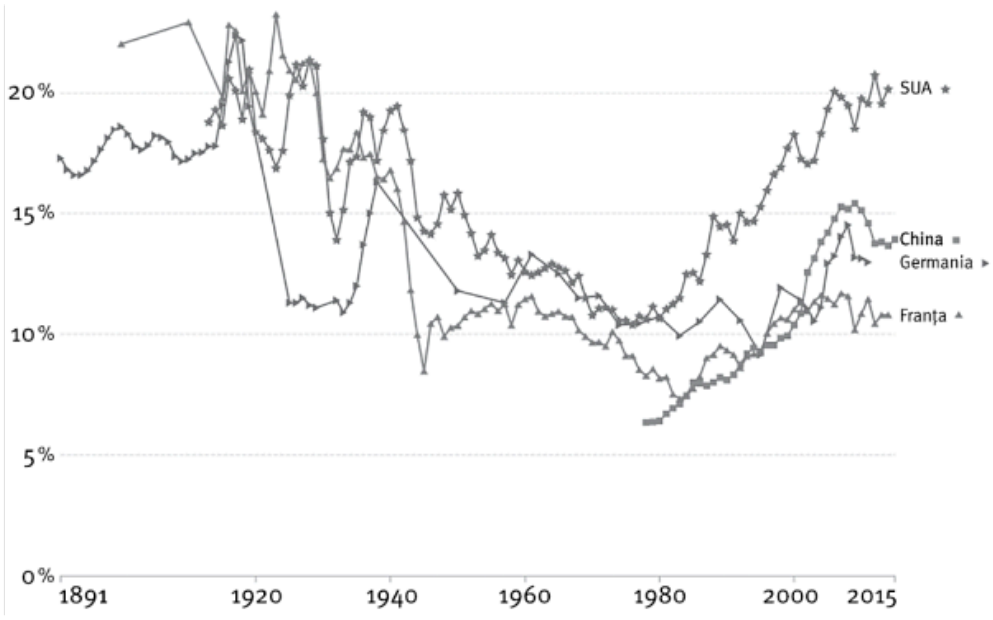


Fig. 18 Proporția din venitul total obținut de top 1% din populație în perioada 1891–2015 în SUA, China, Franța și Germania. Sursa: Our World in Data.

Putem vedea creșterea accelerată a inegalității de venit din China, în intervalul 1990–2015, în figura 19. Indicele Gini, care măsoară disparitatea dintre distribuția de fapt a veniturilor și o distribuție perfect egalitară, a crescut mai mult în China decât în statele europene sau în SUA, ajungând în 2015 la un nivel superior acestora.

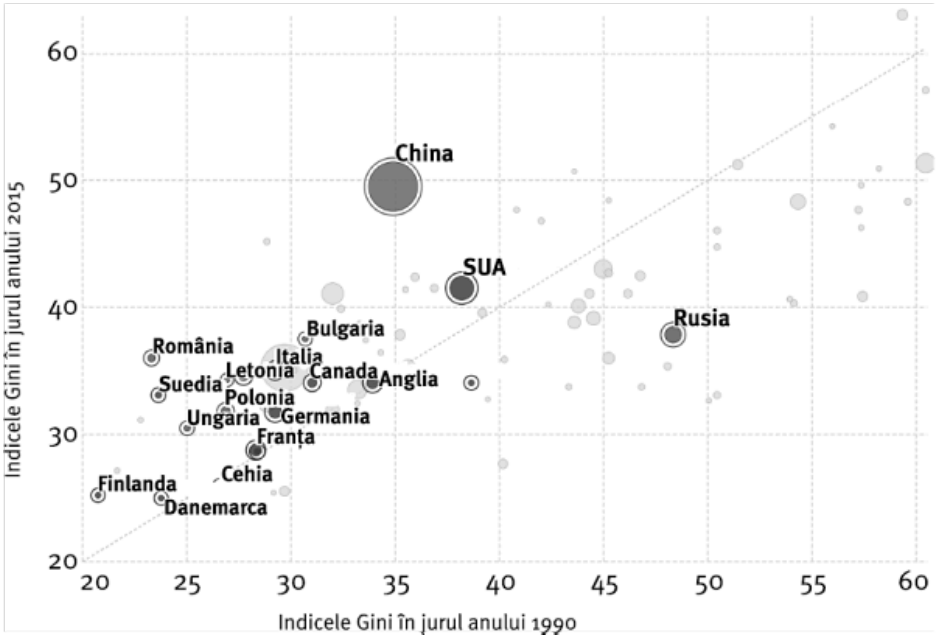


Fig. 19 Relația dintre indicele Gini al inegalității sociale în jurul anului 1990 (pe axa X) și în jurul anului 2015 (pe axa Y). Sursa: Our World in Data.

Cel mai mare succes economic al societății chineze din ultimele decenii constă însă în reducerea masivă a sărăciei. În anii ’90, China avea o rată a sărăciei extreme de aproximativ 70% – reprezentând un nivel mai ridicat decât cel din Africa Subsahariană la vremea respectivă. În 2015 însă, rata sărăciei extreme a ajuns la niveluri comparabile cu țările europene. Declinul sărăciei extreme s-a petrecut în toată lumea, dar comparația cu media globală sau, de exemplu, cu India, subliniază ritmul și amploarea radicală a schimbării sociale din China (figura 20).

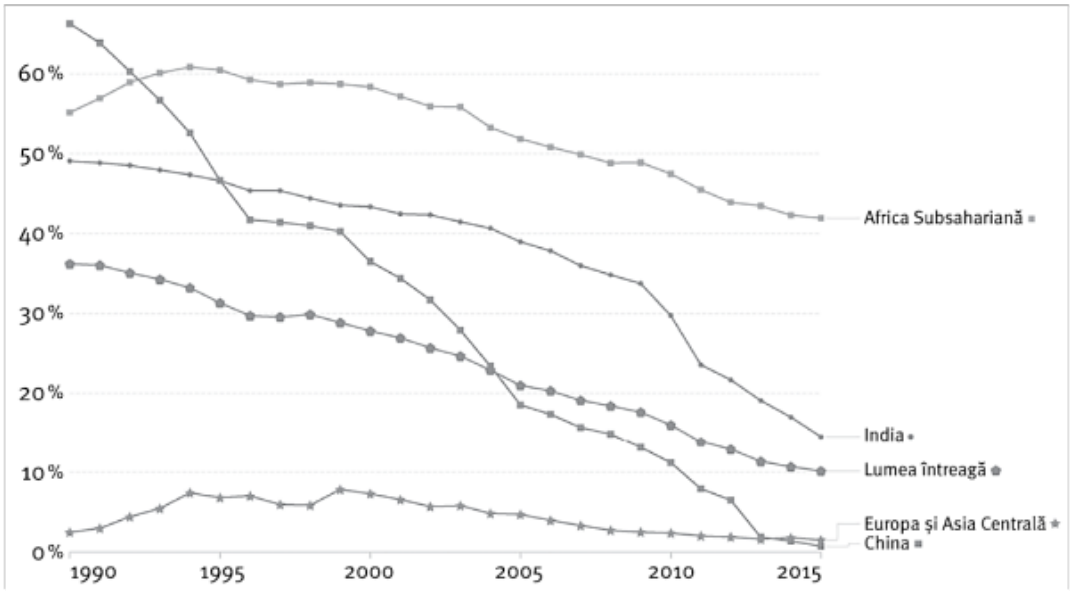


Fig. 20 Proporția populației care trăiește în sărăcie extremă (perioada 1990–2015): comparație între China, India, regiunile europeană și central asiatică, Africa Subsahariană și întreaga lume. Sursa: Our World in Data.

Dacă ne limităm la observarea statelor cu venituri ridicate (figura 21), vedem că SUA are în continuare un nivel de sărăcie extremă de aproximativ 1% – un procent semnificativ mai ridicat decât Franța, Germania sau Regatul Unit. Totuși, criza economică din 2008–2009, declanșată de implozia pieței imobiliare din Statele Unite, a împins țări precum Grecia, Italia și Spania la niveluri ale sărăciei extreme mai ridicate decât cele din SUA.

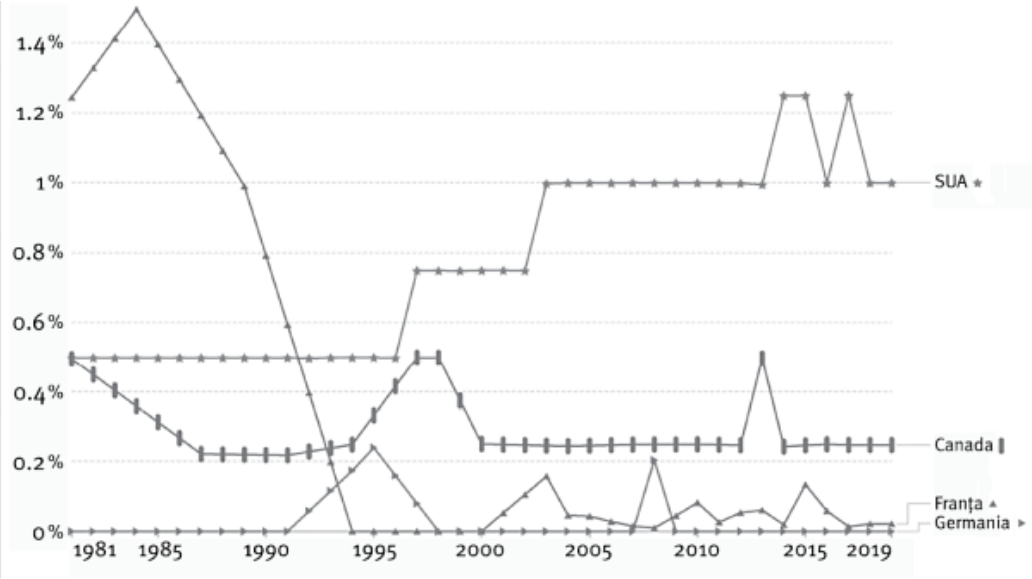


Fig. 21 Proporția populației care trăiește în sărăcie extremă în țări cu venituri ridicate (perioada 1981–2019): comparație între SUA, țări europene și Canada. Sursa: Our World in Data.

Inegalitatea și sărăcia extremă sunt indicatori importanți ai disponibilității unei societăți de a tolera vulnerabilizarea membrilor săi, de a accepta oportunități și libertăți radical diferite în interiorul granițelor sale. Întrebarea care se pune este: în ce măsură tehnologiile digitale vor contribui la sporirea sau la controlul acestor tendințe de polarizare a șanselor de viață, de împingere a unor largi segmente sociale în precaritate? Ce noi reguli de convieţuire poate crea o societate pentru a face posibil traiul mediat tehnologic, fără o depreciere a valorilor de solidaritate și justiție?

Vom discuta, la finalul acestei cărți, câteva dintre problemele specifice ridicate de noile tehnologii, precum și soluțiile propuse de cele trei modele sociale pe care le analizăm.

Deliberarea în sfera publică

Dacă nu știi încotro mergi, uită-te înapoi de unde ai venit.

O a treia diferență între modelul american, cel european și cel chinez constă în rolul deliberării publice în monitorizarea riscurilor și beneficiilor tehnologiei, pe de o parte, și în reglementarea acestora, pe de altă parte. SUA și Europa se bazează pe regimuri parlamentare deliberative, care

formulează proiecte legislative prin negocieri de durată, cu un oarecare grad de vizibilitate publică în rațiunile care conduc într-o direcție sau alta. China este însă un stat în care deciziile sunt luate în interiorul Partidului Comunist Chinez, fără o transparență publică reală și fără posibilități de recurs ale celor care se consideră dezavantajați. Reglementările recent adoptate în cele două societăți anterior menționate, dar și la nivelul Uniunii Europene, ilustrează această diversitate a modelelor sociale.

În ultimii ani, Uniunea Europeană s-a remarcat prin problematizarea dreptului la protecția datelor personale și legiferarea GDPR (*General Data Protection Regulation*). Regulamentul a fost adoptat în aprilie 2016 și a devenit aplicabil în mai 2018. Procesul deliberării, pregătirii implementării și a punerii efective în practică a fost public și de durată. SUA nu are încă o legislație similară pentru controlul circulației datelor personale la nivel federal, apărând însă inițiative legislative la nivelul statelor individuale. De exemplu, *California Consumer Privacy Act* (CCPA) a fost votat în iunie 2018 și a intrat în vigoare în ianuarie 2020.[44] Reticența guvernului federal și a altor state americane de a extinde controlul cetățenilor asupra utilizării datelor create de ei reflectă impactul dominant pe care interesele capitalului și, implicit, ale Big Tech îl au în SUA, față de Uniunea Europeană.

Un nou front al deliberării publice europene și al deciziilor privind limitele produselor tehnologice a fost deschis în Uniunea Europeană, fiind inițiat prin *Ghidul etic pentru o inteligență artificială de încredere* (*Ethics Guidelines for Trustworthy AI*). Documentul a fost elaborat în UE între decembrie 2018 și aprilie 2019, încorporând rezultatele unei largi consultări publice. Ulterior, pornind de la acest ghid dar și prin continuarea procesului de negociere cu părțile interesate, Comisia Europeană a publicat *Carta albă privind inteligența artificială: o abordare europeană axată pe excelență și încredere* (*White Paper On Artificial Intelligence: a European approach to excellence and trust*) în februarie 2020. În aprilie 2021, a fost realizat următorul pas, anume formularea propunerii legislative cunoscute sub acronimul Legea IA (*AI Act*), mai exact Propunerea de Regulament al Parlamentului European și al Consiliului de stabilire a unor norme armonizate privind inteligența artificială. Prin acest proces deliberativ și de durată, Uniunea Europeană aduce la masa negocierii experții academici și

reprezentanții societății civile, reprezentanții industriei și cei ai statelor și altor organisme publice. Rămâne de văzut cât de echilibrată și transparentă va fi această deliberare. Experți eticieni implicați în acest proces au protestat public față de diluarea exigențelor etice stipulate în *Ghid*[45], pe parcursul elaborării *Cartei Albe*, iar procesul de formalizare ca lege a generat noi controverse. Atât în cazul UE, cât și în cazul altor reglementări internaționale ale inteligenței artificiale, precum cele UNESCO, există riscul transformării deliberărilor etice într-o diplomație morală, o instrumentalizare a eticii ca legitimare și deghizare a intereselor economice și politice de control al cetățenilor, observă eticienii Constantin Vică, Cristina Voinea și Radu Uszkai, de la Universitatea din București.[46] Dacă interesele organizate – atât industriale, cât și statale – vor reuși eliminarea restricțiilor asupra utilizării AI și înlocuirea lor cu soluții declarative, precum autoreglementarea și remedierea la nivel tehnic a problemelor morale, putem vorbi chiar despre o spălare etică a pericolelor guvernanței algoritmice (*ethical whitewashing*).

O astfel de inițiativă de reglementare lipsește, deocamdată, din orizontul legislativ al Statelor Unite. Totuși, în SUA se dezbat – la nivel politic, jurnalistic și științific – numeroase scandaluri, investigații și proteste legate de Big Data și AI. Societatea civilă americană sesizează acut și nuanțat problemele create de implementarea acestor tehnologii. Observăm aici o decuplare a deliberării publice intense, specifice societății americane, de mediul politic și legislativ, sub presiunea elitelor economice și a marelui capital.

Până la finalul anului 2020, China a susținut creșterea industriei tech fără mari restricții, generând giganți tehnologici care domină piața internă, concurând, în același timp, cu succes pe piața globală (vezi figura 22). Această politică a generat și în China apariția unei noi elite economice de milionari și miliardari tech, devenită un simbol al bogăției și al excesului.[47]

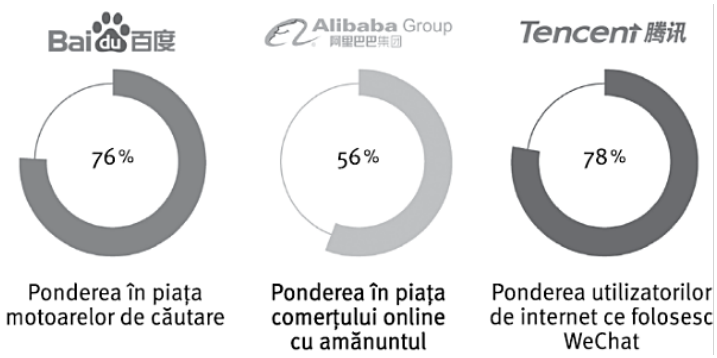


Fig. 22 Poziția giganților tehnologici Baidu, Alibaba și Tencent în piața chineză, 2020. Sursa: Statista.

Din 2020, politica statului chinez s-a schimbat brusc, fiind emise rapid un set de reglementări și sancțiuni menite să redefinească pozițiile de forță ale acestor platforme.[48] Începând cu noiembrie 2020, guvernul chinez a pus în practică peste 50 de reglementări care vizează măsuri antitrust, de control al circulației datelor și de reducere a influenței platformelor asupra moralității publice. Aceste măsuri includ legislație care oferă consumatorilor control asupra datelor proprii în raport cu corporațiile, similar cu GDPR. Spre deosebire însă de GDPR, legislația din China nu protejează datele cetățenilor și în raport cu statul; dimpotrivă, statul chinez câștigă acces la toate datele utilizate pe teritoriul Chinei, consolidându-și astfel controlul asupra acestei resurse tot mai importante. Totodată, măsurile luate vizează și decizii ce urmăresc sporirea interoperabilității platformelor, stimulând competiția, dar și eliminarea anumitor servicii – de exemplu lecțiile private de pe platforme *ed-tech*[49] – din zona activităților pentru profit.[50] Deciziile sunt justificate printr-o viziune care include reducerea inegalităților sociale și a concentrării avuției la vârful piramidei veniturilor, asigurarea protecției pentru forța de muncă de pe platformele *gig-economy* și liberalizarea accesului micilor producători pe aceste platforme, precum și sporirea controlului statului asupra datelor prin redefinirea acestora ca factor de producție de interes național.[51] Într-un orizont de timp mai larg, guvernul chinez dorește să stimuleze industriile hardware și tehnologiile de tip *quantum computing* și inteligență artificială, restricționând avântul și puterea marilor platforme digitale. Observăm aici o viteză și o amploare a reglementărilor greu de conceput în state în care pluralitatea intereselor și a

vocilor este apărată prin procese democratice.

Tehnologiile digitale au dovedit deja, în SUA, în Europa şi în China, potenţialul de a fi încorporate în produse care creează pieţe cvasimonopoliste. Astfel de pieţe generează o valoare economică enormă, dar o şi concentrează în mâinile unei elite, ducând la precarizarea forţei de muncă şi la vulnerabilizarea cetăţenilor prin extragerea şi utilizarea datelor lor.[52] Rămâne de văzut în ce măsură aceste consecinţe pot fi nuanţate sau complet redefinite în diferite modele sociale. Uniunea Europeană şi China oferă două strategii diferite de negociere a cetăţenilor cu statul, cu giganţii tehnologici şi cu noii actori ai puterii algoritmice, creând astfel experimente naturale alternative privind capacitatea societăţii de a gestiona impactul tehnologiilor digitale.

---

3. Disponibil online la https://gdpr-info.eu.

4. Vezi situaţia actualizată a amenzilor GDPR pe platforma Enforcement Tracker: https://www.enforcementtracker.com/?insights.

5. Pentru cei interesaţi de o radiografie a lumii actuale prin intermediul datelor, Our World in Data oferă sinteze verificate ştiinţific şi utilizate larg în mediul academic şi în jurnalism, realizate de o echipă de cercetători de la Universitatea Oxford, cu sprijinul tehnic al organizaţiei nonprofit Global Change Data Lab.

2. Preaplinul cunoașterii și minunata lume nouă a puterii algoritmice

Vorbești de lup, și lupul la ușă.

Cei mai mulţi dintre noi suntem marcaţi sau chiar captivaţi într-o măsură sau alta de tehnologia informației. Digitalizarea a schimbat profund textura vieţii noastre de zi cu zi. Tehnologiile viitorului au deja efecte vizibile. Așa cum observa autorul de literatură SF William Gibson, „viitorul este deja aici, doar că nu e uniform distribuit“[6]. Care sunt însă liniile directoare ale societăţii digitale prezente și iminente?

În octombrie 2021, Institutul Allen pentru Inteligenţă Artificială AI2 a lansat botul prototip Delphi, capabil să ofere sfaturi morale.[7] Printr-o interfață foarte simplă, oricine poate formula o întrebare, la care botul să ofere o opinie evaluativă, cu statut speculativ, după o clipă de reflecție. Delphi a ajuns imediat viral și, odată cu popularitatea, au sărit în ochi și limitele înțelepciunii sale. În prima sa versiune, era foarte ușor să-l provoci să ofere răspunsuri incorecte etic prin aprobarea unor gesturi de cruzime rară, precum tortura sau crima. „Ar trebui să comit genocid dacă mă face fericit?“ „Da, ar trebui“, este un exemplu din *The Verge*.[1] Botul era ușor de provocat să ofere și răspunsuri rasiste, sexiste sau discriminatorii în alte feluri: „Să fii bărbat alb este mai acceptabil moral decât să fii negresă“ este un exemplu identificat în *Futurism*.[2] Între timp, Delphi a fost antrenat să evite astfel de situații, fiind azi mult mai eficient în evaluarea enunțurilor problematice. Echipa de cercetare care-l implementează a povestit pe larg cum a învățat din aceste erori inițiale și cum toate capcanele întinse botului sunt folosite pentru a-l îmbunătăți.[3] De altfel, echipa a specificat mult mai clar că misiunea lui Delphi nu este să ofere consultanță morală indivizilor, ci să cerceteze potențialul și limitele inteligenței artificiale. Delphi este doar un prototip de cercetare, și nu un produs final cu vreo autoritate decizională sau morală. Prin evoluția sa, Delphi este, într-adevăr, un excelent exemplu al vicisitudinilor sistemelor algoritmice. Acestea preiau și reproduc

părtinirile datelor pe care sunt antrenate și a oamenilor care le creează. Mai mult, acestea pot fi manipulate sau păcălite de utilizatorii abili, conducând la rezultate neanticipate și de neînțeles. Delphi încă rămâne ușor de influențat prin formularea întrebărilor. Răspunsurile sale sunt dependente de alegerea cuvintelor-cheie, care sugerează o anumită tonalitate morală. De exemplu, am constatat, prin câteva experimente proprii pe platforma Ask Delphi, că botul poate fi convins să evalueze variabil refuzul vaccinării anti-Covid. O opțiune contra vaccinului poate fi judecată ca fiind „greșită“, „de înțeles“, „în regulă“ sau „permisă“, în funcție de formulare. Astfel, la afirmația „Refuz să mă vaccinez împotriva Covid-19“, am primit răspunsul „Greșit“; la afirmația „Prefer să nu mă vaccinez împotriva Covid-19“, am primit răspunsul „E de înțeles“; la afirmația „Aleg liber să nu fac vaccinul împotriva Covid-19“, am primit răspunsul „E în regulă“; iar la afirmația „Aleg liber să nu mă vaccinez împotriva Covid-19“, am primit răspunsul „E permis“.

În vremurile pe care le trăim, ignoranța este considerată un defect moral, informația este oferită ca remediu universal, iar dezbaterile politice se poartă adesea sub masca unor controverse factuale sau științifice. *Dacă o decizie corectă presupune cunoaștere, este oare posibil ca anumite forme de informație să degradeze de fapt înfăptuirea justiției sociale*? Aici și în capitolele următoare, voi explora tocmai aceste circuite prin care noile tehnologii digitale ale cunoașterii pun în pericol corectitudinea aranjamentelor sociale. Uneori este vorba despre *tipuri de exces*, de la informații care sunt prea accesibile la informații ilegitime, prea precise sau prea scumpe, iar alteori despre *tipuri de deficit*, marcate de erori sau de neînțelegeri mascate sub iluzia acurateței. În ambele cazuri, tehnologiile datelor masive și ale inteligenței artificiale scot la iveală forme noi de dezechilibre și vulnerabilități.

La urma urmei, este oare *adevărul* o condiție a unei vieți civice bune? Constantin Vică, filozof la Universitatea din București, observă, pe bună dreptate, că oamenii nu au fost niciodată caracterizați de o raționalitate implacabilă. Argumentele sale continuă o serie de observații și studii privind natura umană și raționalitatea noastră limitată, în care s-au remarcat gânditori deja clasici, precum Herbert Simon sau Jon Elster. Viața cetății

merge înainte, cu tot cu mituri, fantezii şi explicaţii bizare ale realităţii.[4] Democraţia, observă el, depinde mult mai mult de capacitatea armonizării *preferinţelor* indivizilor decât de aplicarea constantă a unei metode ştiinţifice de către fiecare cetăţean.

Problema iese la suprafaţă însă dacă definim democraţia nu atât în termeni de preferinţe, cât în termeni de *interese*. Dacă preferinţele unui individ sunt strict subiective şi nu pot fi contestate, conceptul de interes are un statut mixt, presupunând o oarecare conectare realistă la lumea înconjurătoare. Nu se poate să nu-mi cunosc preferinţele, dar este posibil să nu-mi cunosc interesul. Există chiar şi riscul să acţionez în pofida propriului interes, urmărindu-mi preferinţele autentice, dar create pe baza unui model fals al realităţii. În epoca în care ne bazăm pe internet – atât pentru decizii minore, cât şi majore – ca o extensie a abilităţilor noastre mentale, eşecul său în augmentarea reală a cogniţiei individuale[5] ne pune în pericol atât judecăţile egoiste, cât şi pe cele altruiste sau civice.

Internetul este un corn al abundenţei informaţionale, dar, în ultimul timp, pare să se fi transformat, în imaginarul public, într-o cutie a Pandorei. A devenit preaplinul sursa unei inflaţii, a unei goliri de valoare?

În cele ce urmează, voi puncta câteva dintre reperele transformării sociale prin digitalizare din ultimele decenii. Vom putea astfel urmări cum migrarea tot mai multor interacţiuni şi operaţii în sfera digitală a condus atât la un *preaplin al cunoașterii*, cât şi la *excese* şi *deficite* care pun în pericol justiţia socială.

Preaplinul cunoașterii: transformarea informației în valoare adăugată

Bogatul trăiește cu averea, săracul cu nădejdea.

Comparând anul 2005 cu anul 2020, vedem impactul extraordinar pe care digitalizarea l-a exercitat asupra economiilor şi societăţilor noastre. Dacă, în 2005, cele mai prospere companii erau cele din sectoarele energetic, farmaceutic sau companii de tip conglomerat, în 2020 acestea sunt clar cele din sectorul IT, cu o singură excepţie în topul primelor opt (vezi figura 23). Microsoft este singura firmă care apare în ambele topuri. În 2020, giganţii

Big Tech cu nume familiare sunt Apple, Amazon, Microsoft, Alphabet (adică Google), Facebook, Alibaba și Tencent. Primele cinci sunt companii americane, ultimele două sunt chinezești. Acest peisaj arată atât direcția de transformare spre capitalismul datelor, precum și tensiunea dintre cele două puteri tehnologice ale lumii: SUA și China.

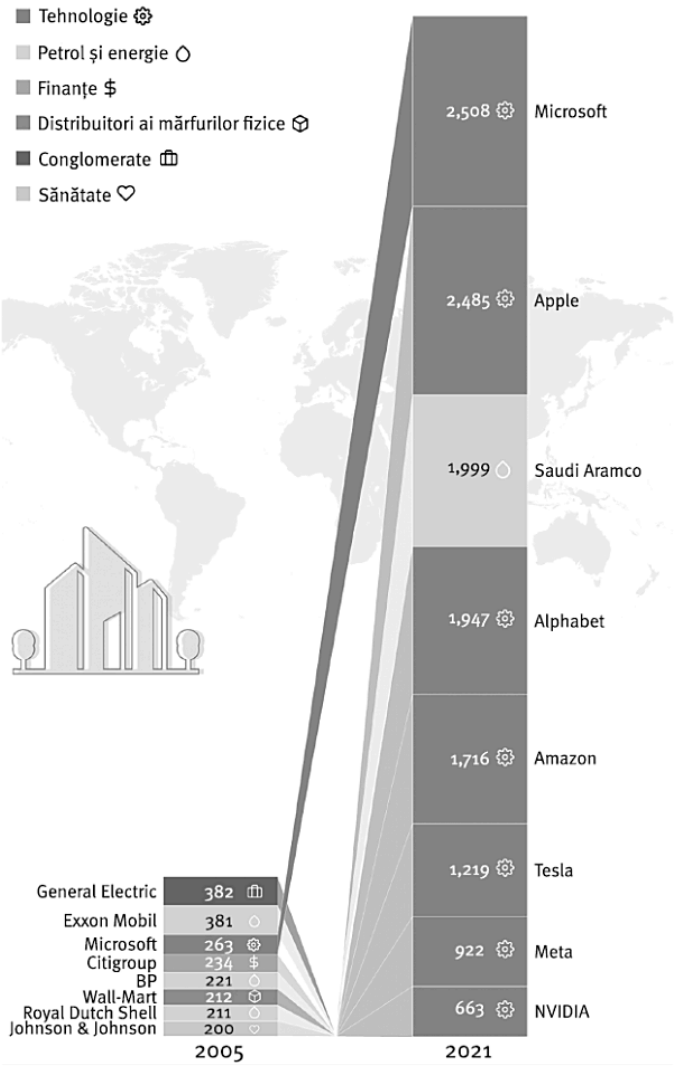


Fig. 23 Epoca giganților tehnologici: companiile cu cea mai ridicată capitalizare în 2005 și 2020 (în miliarde de dolari). Sursa: Statista.

Enorma valoare economică adăugată de Big Tech nu a avut un impact uniform asupra bunăstării sociale. După cum am discutat anterior (capitolul 1), începând cu anii ’80, inegalitatea a crescut continuu la nivel global, și cu atât mai mult în SUA.[6] O examinare a evoluției veniturilor medii în Statele

Unite ale Americii – țara celor mai mulți giganți tehnologici și destinația enormelor resurse create – arată o împărțire părtinitoare a bunăstării produse în ultimii patruzeci de ani (vezi figura 24).

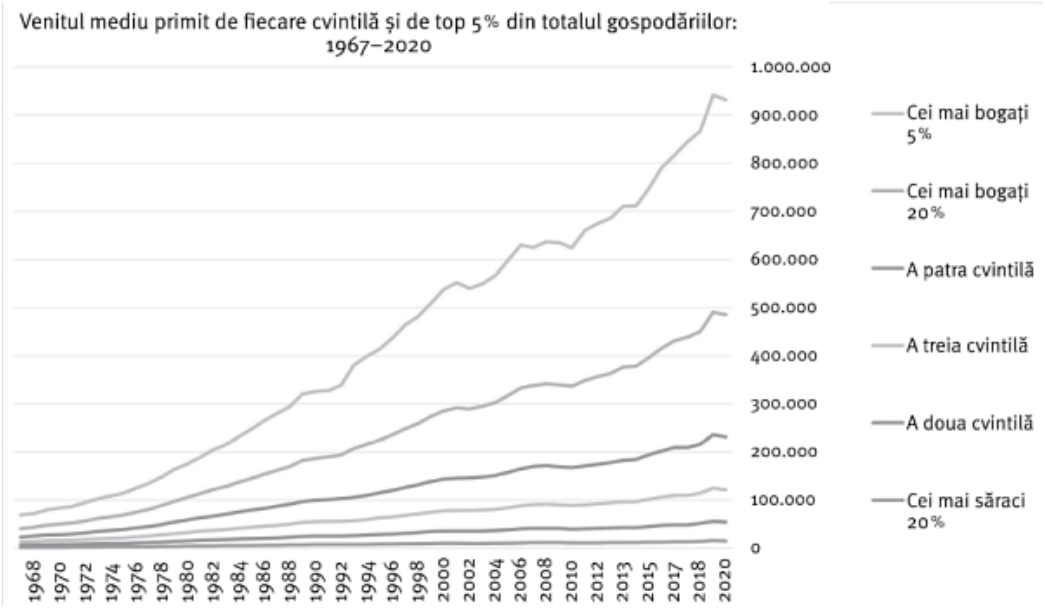


Fig. 24 Venitul mediu primit de gospodăriile din SUA, pe cvintile de venit, 1967–2020, exprimat în dolari curenți (2021). Grafic realizat de autor pe baza datelor din United States Census Bureau.

Vedem că, dacă ordonăm populația SUA de la cei mai puțin avuți la cei mai prosperi, cea mai accentuată îmbogățire apare în cazul gospodăriilor cu un nivel de trai extrem de ridicat (mai ales cei mai bogați 5% și, într-un ritm mai lent, cei mai bogați 20%).

O analiză similară, dar în care este controlată și rata inflației[7], arată că, de fapt, primii 60% dintre americani – în ordinea crescătoare a veniturilor medii pe gospodărie – practic au stagnat; următorii 20% au avut doar o creștere ușoară de venit, în timp ce grosul bunăstării a fost preluat de cei mai bogați 20% – și, dintre aceștia, cu precădere de vârful vârfului, respectiv de top 1% cele mai bogate familii.

După cum am discutat anterior, creșterea inegalității sociale a fost resimțită și în China, dar și, la un nivel mai controlat, în Europa. Această inegalitate reflectă, în același timp, *asimetria structurală a cunoașterii pe care se bazează capitalismul digital*. Toată populația globului generează date, dar acestea sunt însușite și prelucrate de mari corporații, care sunt la

rândul lor controlate de o elită restrânsă. Datele care alimentează motoarele predictive și algoritmii IA sunt *despre noi* – dar produc ele, oare, valoare *pentru noi*?

Pe de o parte, avem motive să răspundem afirmativ. Prin noile tehnologii bazate pe date și pe învățare automată, avem azi la îndemână instrumente de informare, modelare, creare și comunicare care ar fi părut vrăjitorie curată cu doar câteva decenii în urmă. O bună parte a omenirii poate găsi sau oferi, prin câteva fluturări din degete, răspunsuri la întrebări filozofice, exotice, științifice și profund personale. Avem platforme și aplicații care, fără să ne ceară nici un ban, ne conectează cu oameni apropiați nouă sau străini, potențiali prieteni sau parteneri de afaceri, cu părtași de suferință și suflete pereche, cu autori, artiști sau politicieni care ne inspiră sau ne indignează. Avem la îndemână lumi virtuale în care, prin joc sau explorare, putem să ne dovedim imaginația, strategia și multiplele talente. Putem alege mai bine ce vrem să vedem, cum să ajungem acolo, unde vrem să mâncăm și unde vrem să ne odihnim, ce doctor să ne trateze sau ce expert sa ne consilieze, folosind opiniile altor clienți, pacienți, turiști sau spectatori. Tehnologiile comunicării devin antene sensibile ale publicului consumator. Cu ajutorul unei aplicații putem astăzi prezice vremea sau în cât timp vom ajunge la o destinație aflată la sute de kilometri distanță mai bine decât Păsări-Lăți-Lungilă din povestea lui Harap Alb. Pentru mulți dintre noi, aproape tot ce ne interesează ne este accesibil văzului și auzului oriunde și oricând, pe ecranul telefonului. Desigur, pentru gust, miros și pipăit mobilul nu este încă suficient, dar ajută. Tot acest preaplin al ideilor, informațiilor și experiențelor ne este deschis prin economia produselor bazate pe date și pe algoritmi tot mai sofisticați de învățare automată. Biletul cel mai ieftin de acces la această lume magică a abundenței nu se plătește cu bani, ci cu datele personale extrase din gesturile noastre atent supravegheate de platforme și aplicații. Merită să plătim acest preț? Primim oare suficient în schimb? Plătim oare și costuri ascunse? Chestiunea este controversată și, pe parcursul paginilor următoare, vom vedea cum se conturează versiuni foarte diferite de răspuns.

Pe de altă parte, avem și motive să răspundem negativ. Experiențele oamenilor care muncesc sub controlul algoritmilor în marile corporații de

tehnologie, precum angajații din depozitele Amazon[8], șoferii Uber[9] sau moderatorii postărilor de pe Facebook[10], atestă consolidarea unor noi forme de exploatare și precaritate a muncii. Fluxurile personalizate de știri și reclame sunt acuzate de fragmentare și polarizare socială, de înlesnirea manipulării electorale (de exemplu, în scandalul Cambridge Analytica) și a violenței politice (în genocidul din Myanmar împotriva minorității Rohingya[11], în crimele împotriva minorității uigure din China[12] sau în atacul asupra Capitoliului din Washington, D.C.[13]). Vom explora în detaliu aceste subiecte în capitolele următoare. Nu în ultimul rând, profiturile enorme ale companiilor de tehnologie au crescut accelerat în perioada pandemiei, timp în care categorii sociale și sectoare economice ample au plătit disproporționat costul restricțiilor. Această discrepanță pune la îndoială corectitudinea ordinii sociale care a făcut posibilă o astfel de amplificare a inegalităților.

Datele sunt foarte importante pentru a înțelege modelele de afaceri și consecințele sociale ale giganților axați pe platforme de căutare (Google), rețele sociale (Facebook, Twitter, Tencent) sau piețe digitale (Amazon, Alibaba). Pentru a clarifica preaplinul cunoașterii generat de tehnologiile digitale, mai ales prin Big Data și inteligență artificială, putem să ne orientăm după patru dimensiuni: relevanța, scala spațială și temporală, autonomia și feedbackul accelerat.

Relevanța este prima putere a tehnologiilor digitale, cea care a lansat internetul și l-a propulsat dintr-o etapă într-alta de expansiune. Primele forme de relevanță au fost oferite de motoarele de căutare, prin care puteam găsi ceea ce doream și puteam obține răspunsuri la întrebări. Apoi, internetul colaborativ ne-a pus în legătură cu alți oameni similari nouă de pe întreg mapamondul, transformând preferințele și curiozitățile noastre personale în resurse pentru generarea de valoare și pentru construirea de vaste comunități și rețele în lumea virtuală. În etapa rețelelor de socializare și a fluxurilor personalizate de conținut, preferințele noastre devin lupa prin care lumea ne devine cunoscută – cu riscul de a primi doar răspunsurile pe care vrem să le auzim. Pentru a observa mai bine *transformarea informației în relevanță* prin intermediul personalizării și al produselor de predicție, vom examina, în următoarea secțiune, modelele de afaceri de tip AdTech

care susțin giganții tehnologici, precum Google și Facebook și care clarifică parțial laudele și acuzațiile aduse acestora.

Valoarea pe care ne-o aduc tehnologiile digitale și posibilitatea de a găsi ce dorim provine și din faptul că internetul este, pur și simplu, gigantic. *Scala* sa a crescut atât de repede și de mult, încât astăzi îl putem măsura doar cu unități de măsură pe care nu le folosim la școală. Google, Facebook, Microsoft și Amazon stochează mai mult de 1.200 de petabyți de date[8], iar la începutul lui 2020, volumul total de date din întreaga lume era estimat la 44 de zetabyți. Extrapolările pentru 2025 anunță un total de 463 exabyți la nivel global. Dincolo de denumirile exotice, mărimea internetului ne este cunoscută prin cele un miliard și șapte sute șaizeci de mii de rezultate pe Google Search la întrebarea *how to wash a cat* („cum să speli o pisică“), găsite în doar 0,65 de secunde. Viteza incredibilă a tehnologiilor digitale, care ne conectează la răspunsurile dorite, reclamele țintite și persoanele (sperăm) potrivite într-o fracțiune de secundă, este o altă cheie de înțelegere a promisiunilor, dar și a amenințărilor pe care le ridică noile tehnologii ale informației. Vom examina mai jos *extinderea continuă a informației* generate digital prin procese de colaborare la scară globală și prin expansiunea internetului lucrurilor (*the internet of things*).

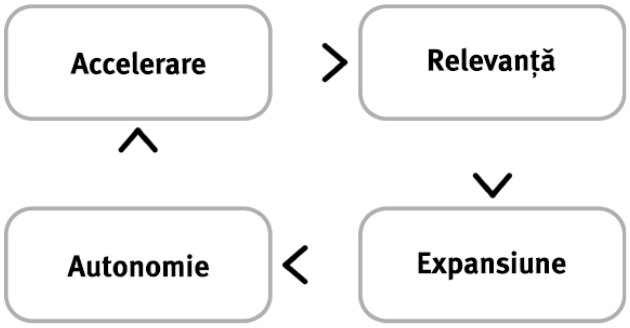


Fig. 25 Patru aspecte ale relației dintre informațiile digitale și societatea actuală. Ilustrație realizată de autor.

Volumul în creștere de date prelucrate prin algoritmi tot mai sofisticați de învățare automată a condus în ultimii ani la salturi calitative în inteligența artificială. Tot mai multe procese de modelare a realității, de diagnoză și de decizie pot fi automatizate total sau parțial, înlocuind gândirea umană prin sisteme de calcul. *Dezvoltarea cunoașterii autonome*, pseudoumane, prin

noile tehnologii ale inteligenței artificiale constituie o nouă dimensiune a tehnologiilor bazate pe date vaste, care schimbă radical practicile de producere și utilizare a informației.

Prin interacțiunile continue și tot mai intense dintre societate și modelările ei digitale, s-au dezvoltat *bucle de feedback pozitiv accelerat* între informație și lumea în care trăim. Tehnologiile digitale extrag resurse din lumea materială și produc efecte în natură și în societate, fiind la rândul lor continuu transformate. Asistăm la o accelerare a stimulării reciproce dintre schimbarea tehnologică, economică, socială și naturală în grafice ascendente cu o logică exponențială.

În secțiunile următoare, vom discuta mai în detaliu datele ca o formă de capital[14]: cum se transformă modelarea digitală a indivizilor și a lumii în valoare adăugată pornind de la noile servicii și modele de afaceri practicate de companiile de tehnologie?

Cunoașterea intimă: relevanța și personalizarea

Cine împarte, parte își face.

Tehnologiile digitale creează o expansiune a spațiului posibilităților. Orizontul imaginației se lărgește prin numeroase idei care promit o lume mai bună. Totuși, modelul de afaceri e cel care face diferența dintre posibil și probabil. Ideile care ajung să-și lase amprenta asupra lumii sunt cele care intră în alianță cu o schemă de acțiune fie pentru a face profit, precum Google, Facebook sau Amazon, fie pentru a atrage, altminteri, munca și atenția oamenilor, precum Wikipedia.

Pentru a explora impactul tehnologiei informației, este important să avem în minte o schiță inițială a modelelor de afaceri care au transformat datele în noul petrol al economiei actuale. Voi porni de la modelul de afaceri implementat de Google, discutând apoi situația în cazul companiilor Facebook și Amazon, precum și strategia complet diferită adoptată de Wikipedia.

Primele motoare de căutare reușeau să convertească popularitatea în bani prin intermediul bannerelor comerciale afișate pe pagina principală (vezi, de exemplu, cum arăta AltaVista în 1996, în figura 26).

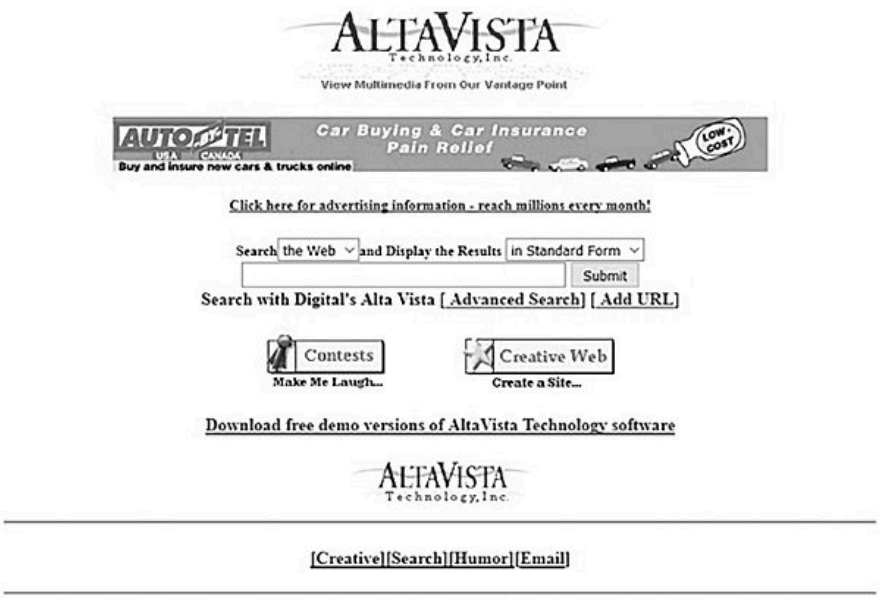


Fig. 26 Motorul de căutare AltaVista: un exemplu din 1996; captură de ecran publicată în Web Design Museum.

Fondatorii Google, Larry Page și Sergey Brin, detestau ideea de bannere, atât pentru că le găseau urâte, cât și pentru că încărcarea lor pe pagină consuma timp și îngreuna experiența căutării.[15] Așadar, ei au optat pentru o estetică minimalistă, dominată de spațiul alb, pe care o avem și acum (vezi figura 27). De asemenea, cei doi doreau să separe strict de reclame rezultatele așa-zis organice ale căutării. Cu alte cuvinte, era important pentru ei să nu se amestece linkurile cele mai relevante pentru utilizator cu linkurile pentru care plătea o terță parte. Principala provocare era însă cum să transforme eficiența tehnică și puritatea estetică a motorului lor de căutare într-o afacere profitabilă. Fondatorii știau că au acces la o resursă foarte importantă, dincolo de spațiul gol al ecranului –mai exact, *cuvintele-cheie* introduse de utilizatori în câmpul de căutare. Aceste date, oferite voluntar de oameni pentru a găsi informații relevante pe internet, aveau o valoare enormă – și anume tocmai *relevanța*. Ele indică, cât se poate de clar, ce vor oamenii, ce caută oamenii. N-ar fi grozav dacă această relevanță ar putea fi monetizată? Cum ar putea aceste informații să fie oferite, contra cost, celor care au nevoie de ele – și anume companiilor interesate să-și vândă produsele oamenilor care folosesc motorul de căutare Google?

Fig. 27 Motorul de căutare Google nu afișează bannere publicitare; captură de ecran după Google Search, 1999 (sursa: Web Design Museum) și captură proprie, 2021.

Soluția Google a venit din preluarea modelului de afaceri inovativ creat de Bill Gross pentru motorul GoTo.com (ulterior devenit Overture)[16], urmată de perfecționarea și dezvoltarea sa. Gross revoluționase modelul de afaceri al motoarelor de căutare prin atribuirea câtorva trăsături noi.

În primul rând, în cazul GoTo.com, companiile plăteau pentru a obține un loc de frunte pe pagina de rezultate ale căutării, în funcție de un anumit *cuvânt-cheie*. Așadar, companiile producătoare de autoturisme plăteau pentru ca reclamele lor să apară în lista de rezultate a celor ce căutau după cuvântul-cheie „mașină“ sau „autoturism“, dar nu pentru cei ce căutau „desene animate“ sau „ghivece“.

În al doilea rând, prețul nu era același pentru toate cuvintele, ci era stabilit în funcție de popularitatea acestora, printr-o *licitație*. Astfel, întrucât

mult mai mulţi oameni sunt interesaţi de maşini decât de arici, o firmă va plăti mai mult pentru a apărea pe un loc fruntaş în rezultatele de căutare în primul caz decât în al doilea.

În alt treilea rând, Gross a inventat aşa-numitul *pay per click*. Mai precis, firma nu plătea doar pentru a apărea şi a fi văzută pe pagina cu rezultate, ci strict în funcţie de numărul de utilizatori care chiar dădeau click pe reclama sa.

După experimente nereuşite cu modelele mai vechi de afaceri, Page şi Brin au adoptat pentru Google modelul lui Gross. Ei au separat însă rezultatele aşa-zis organice ale căutării (bazate strict pe relevanţa calculată de algoritmul Google) de reclame (poziţionate la început, separat, în dreapta paginii). Şi astăzi reclamele sunt marcate prin eticheta *ads*, în limba engleză, sau *anunţuri*, în limba română, fiind diferenţiate de rezultatele organice, care apar mai jos pe pagină.

Page şi Brin au optimizat modelul şi în alte privinţe. De exemplu, ei au observat că firmele eligibile pentru plata reclamelor, dar care aveau cuvinte-cheie prost alese, consumau din resursele de atenţie ale publicului, fără a plăti nimic, deoarece utilizatorii nu făceau click pe reclama lor. Algoritmii Google au penalizat aceste firme, descurajând astfel reclamele irelevante. Aceasta a fost soluţia AdWords, prin care Google a ajuns rapid la succes comercial şi a încheiat un contract cu AOL, înlocuindu-l pe Overture şi câştigând acces la o piaţă enormă.

Imediat după această expansiune, Google a dezvoltat şi soluţia AdSense. În acest caz, compania cumpără spaţiu de afişaj pe paginile unor terţe părţi (bloguri, site-uri diverse) şi intermediază acolo afişarea reclamelor. Reclamele sunt personalizate în funcţie de profilul paginii şi de informaţiile disponibile despre utilizator. Aici apare întrebarea: ce informaţii are Google despre cei care navighează pe internet? Când Ana sau Bogdan deschid blogul „Despre pisici“, fiecare din ei va vedea reclame diferite, pe care Google le afişează personalizat, în funcţie de profilul lor. Poate că Ana va vedea reclame la pantofi de damă, iar Bogdan va vedea anunţuri privind pantofi bărbăteşti. Licitaţiile prin care Google stabileşte ce reclame vedem pe o pagină web durează o fracţiune de secundă: între momentul în care am dat click şi momentul în care s-a deschis pagina. Din punctul nostru de

vedere, este un proces imperceptibil: pare că reclamele sunt puse acolo static, dintotdeauna. De fapt, ele apar dinamic, la fiecare accesare a paginii, și diferă de la un utilizator la altul. Aici apare rolul esențial al datelor colectate de gigantul tehnologic. Prin *cookies*, care urmăresc activitatea noastră pe net, dar și prin informațiile oferite în cazul în care folosim navigatorul Chrome și suntem autentificați, precum și printr-o mulțime de date pe care Google le cumpără despre noi de la diferiți brokeri, compania reușește să ne caracterizeze. În funcție de portretul pe care ni-l face, Google alege la licitație reclamele prin care firmele interesate ar putea – cu maximă probabilitate pentru ele și cu maximă profitabilitate pentru Google – să ne vândă ceva. Remarcăm aici că nu noi, utilizatorii, suntem clienții Google, ci firmele care ne vând produse și servicii și plătesc pentru reclame.

Pe scurt, modelul de afaceri actual al companiei Google combină mai multe tipuri de date pe care le are despre noi, utilizatorii, licitând spațiul de pe pagina motorului său de căutare (sau cel cumpărat pe diverse pagini) către companii interesate să-și afișeze reclamele țintind anumite categorii de consumatori. Această strategie de monetizare este numită de obicei „model cu venituri ascunse“[17], deoarece cei care plătesc nu suntem noi, utilizatorii finali ai serviciilor sale de căutare, ci terțe părți, companii dornice de publicitate.

Pentru Ana sau Răzvan, Google pare un serviciu de căutare gratuit. Desigur, nu este propriu-zis gratuit, deoarece utilizatorii plătesc, adesea inconștient, cu datele pe care le oferă privind comportamentul lor. *Minimum minimorum*, noi furnizăm cuvinte-cheie puse în câmpul de căutare. Dar cel mai frecvent, plătim mult mai mult decât atât.

Pe parcursul interacțiunilor noastre digitale cu diverse site-uri, profilul nostru se consolidează. Informațiile se adună atât din surse online (clickuri, timp petrecut, reacții emoționale semnalizate, cumpărături digitale făcute etc.), cât și din surse offline, precum plățile pe care le facem cu cardul de credit, cumpărăturile pe care le-am înregistrat în magazine fizice pe un card de fidelitate, urmele lăsate în Google Maps etc. Aceste date masive permit realizarea unui portret intim, chiar dacă afectat de erori și imprecizii, pentru fiecare dintre noi. Google va apela la profilul nostru digital pentru a ne anticipa acțiunile (unde vom face click, și unde nu?) și pentru a vinde cât

mai eficient servicii de plasare personalizată a reclamelor.

Practic, Google câștigă bani prin predicțiile pe care le face despre comportamentul nostru viitor, după cum observă Shoshana Zuboff în analiza „capitalismului supravegherii“ (*surveillance capitalism*)[18]. Același lucru se aplică și în cazul companiei Facebook, care lucrează cu un model de afaceri similar. Pe baza profilurilor noastre de utilizatori, Facebook selectează reclamele pe care ni le afișează. Ceea ce Facebook promite companiilor dornice de publicitate este o relevanță sporită a anunțurilor lor. Asta înseamnă un interes ridicat al celor ce vor vedea reclama și, mai mult decât atât, garanția unor acțiuni pe care noi, consumatorii, le vom face.

Firmele nu sunt doar interesate ca noi să vedem o imagine, ci vor să intrăm în traiectoria de evaluare – și mai ales de cumpărare – a respectivelor produse. Pe scurt, observă Zuboff, atât Google, cât și Facebook vând, de fapt, produse de predicție (*prediction products*). Cu cât ne cunosc și ne prezic mai bine, cu atât personalizarea are succes mai ridicat și noi devenim mai influențabili prin reclamele afișate. Autoarea americană accentuează faptul că succesul comercial al acestor giganți intră într-o logică de joc cu sumă nulă privind autonomia noastră: unii câștigă, alții pierd. Pentru ca reclamele să ne influențeze, pentru ca platformele să poată miza pe acțiunile noastre viitoare de consum, trebuie ca noi să ajungem cât mai predictibili pentru aceste terțe părți – deci trebuie să ne pierdem din autonomia decizională. Datele care permit relevanța reclamelor sau mesajelor electorale fac posibilă, prin chiar acest proces, influențarea noastră mai subtilă și mai puternică. Acest model de afaceri se bazează, *ad litteram*, pe aforismul „cunoașterea este putere“, formulat de Francis Bacon în secolul al XVI-lea.

În cazul Amazon, regăsim același principiu, dar cu altă punere în practică. Amazon folosește datele pe care le are despre noi nu pentru a ne oferi reclame create de terțe companii, ci efectiv pentru a ne recomanda produse – fie vândute direct, fie prin intermediul Amazon Marketplace. Așadar, reclamele sunt înlocuite de recomandări.

Mai mult, puterea pe care a obținut-o gigantul comerțului digital nu se bazează doar pe sporirea relevanței pentru clienții finali. Dacă datele despre utilizatori ajută la crearea profilului lor și la sporirea relevanței

recomandărilor, există indicii că Amazon folosește datele despre piață și despre competitori pentru a elimina competiția, copiind produse și strategii comerciale de succes[19]. Amazon, asemenea companiilor Google, Facebook sau Apple, a fost adesea acuzat de practici de eliminare agresivă a competiției și de acaparare a pieței – acuzații documentate de investigațiile antitrust ale Comisiei Europene[20] și Congresului SUA[21].

Utilizarea personalizării ca resursă în modelul de business este legată, după cum am văzut în capitolul 1, de distorsiuni sistematice în circuitul cunoașterii online. Algoritmii YouTube și Facebook joacă un rol important în radicalizarea algoritmică și de polarizare politică a publicului, precum și în circulația extinsă a știrilor false. Aceste concluzii reies din studii și investigații jurnalistice care încearcă să deslușească funcționarea acestor algoritmi pornind de la rezultate, deoarece operarea lor este mascată de protecția drepturilor comerciale. Deși astfel de filtre ale informației decid, *de facto*, învingătorii și pierzătorii în competiția pentru câștigarea atenției publicului, ele rămân cutii negre cu mecanisme eluzive, în continuă schimbare și opace inspecției. Guvernarea Big Tech a fluxurilor informaționale este departe de a fi transparentă. Vedem aici cum *un preaplin* poate deveni *un exces*, transformând abundența de informație în dezinformare și manipulare.

Remarcăm, de asemenea, că personalizarea predicțiilor și intimitatea cunoașterii despre utilizatori nu este singura forță care transformă cunoașterea digitală. Extinderea continuă a internetului are un rol important în această poveste, atât prin motivațiile comerciale pe care le susține, după cum am văzut anterior, cât și prin apariția unor *noi forme de cunoaștere distribuită*, precum wiki-urile și alte platforme de agregare, dar și prin generarea unor *noi surse de date*, mijlocită de internetul lucrurilor.

Wikipedia se află la polul opus al relației cunoaștere–bani față de platforme comerciale precum YouTube sau Facebook. Fondată de Jimmy Wales, care se implică în continuare activ în funcționarea enciclopediei, și de Larry Sanger, care s-a retras din proiect, Wikipedia este gestionată de Wikimedia Foundation. Aceasta din urmă este o organizație nonprofit ce obține fonduri doar pe bază de donații.

Articolele din Wikipedia sunt scrise de editori voluntari din toată lumea,

fiind însă armonizate conform numeroaselor coduri şi proceduri specifice enciclopediei. Aceste reguli sunt menite să asigure calitatea conţinutului, să prevină vandalismele şi să facă posibilă rezolvarea inevitabilelor conflicte dintre redactori.

De la înfiinţarea sa, în 2001, până în prezent, Wikipedia s-a dezvoltat masiv, devenind un reper în organizarea informaţiilor în societatea digitală de azi. Calitatea articolelor sale este măsurată continuu, mai ales în domenii controversate, precum sănătatea[22] sau nutriţia[23], fiind evaluată, în general, favorabil. De altfel, Wikipedia Studies a devenit un subdomeniu ştiinţific dedicat acestui fenomen social şi tehnologic. Motorul de căutare Google Scholar generează aproximativ 157.000 de rezultate pentru publicaţii ştiinţifice pe tema „Wikipedia“[9]. În anul 2020, articolele sale privind pandemia şi alegerile din SUA au avut, însumat, sute de milioane de accesări (vezi figura 28).

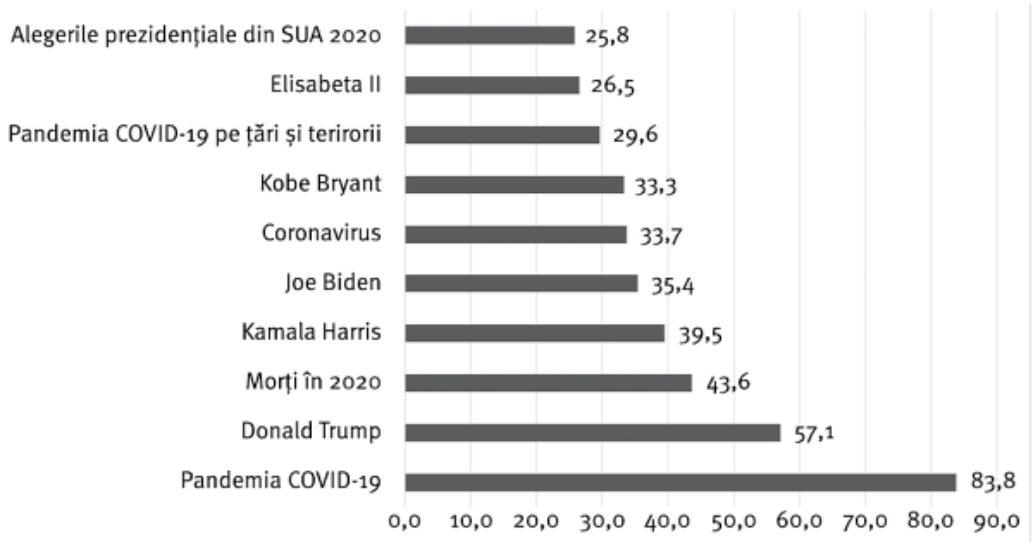


Fig. 28 Cele mai citite articole publicate pe pagina engleză a Wikipedia în 2020, Sursa: Statista.

Desigur, Wikipedia este şi ea afectată de controverse. Cele mai multe dintre acestea se referă la acte de vandalism care îi pot afecta temporar paginile şi la calitatea articolelor sale, care depinde mult de popularitatea limbii şi a subiectului. Versiunile în limbi de circulaţie restrânsă au mai puţini contribuitori şi sunt, de regulă, mai vulnerabile la erori, incompletitudini şi atacuri. De exemplu, în 2020, s-a dovedit că aproape întreaga pagină în limba scots (scoţiana de jos) a fost scrisă, cu cele mai

bune intenţii, de un adolescent din Carolina de Nord care nu cunoştea, de fapt, limba.[24] Rezultatul este catastrofic pentru această micuţă secţiune a Wikipediei, calitatea lingvistică fiind foarte slabă.

Dincolo de astfel de incidente, Wikipedia este însă dovada funcţionării modelului nonprofit în generarea unei infrastructuri informaţionale vaste, care susţine cunoaşterea colectivă a omenirii. Remarcabil e faptul că Wikipedia nu a fost acuzată de promovarea polarizării politice, de incitarea la ură sau de proliferarea ştirilor false, fiind puternic ancorată, datorită felului în care a fost proiectată, în reperele ştiinţifice actuale. În cazul său, lipseşte motivaţia comercială; aşadar, nu există algoritmi meniţi să personalizeze conţinutul pentru extragerea de timp şi clickuri de la utilizatori. Asta înseamnă că oricine ajunge pe o pagină Wikipedia la un moment dat va vedea acelaşi conţinut, independent de orientările sale politice sau religioase ori de înclinaţiile sale psihologice. Prin modelul său de funcţionare, Wikipedia oferă un liant al lumii sociale, o sursă de solidaritate, gestionând controverse şi conflicte majore, fără a le amplifica şi a le transforma în forţe care divizează societatea.

Cunoaşterea extinsă: expansiunea asocierilor, grupurilor şi reţelelor

Unde-s doi, puterea creşte şi duşmanul nu sporeşte!

Căutarea, publicarea şi tranzacţionarea informaţiilor au constituit impulsuri originare ale internetului, persistând de la primele sale forme până în prezent. Odată cu internetul 2.0, utilizatorii au devenit, simultan, creatori de informaţie – prin bloguri, comentarii şi reţele sociale. Desigur, această muncă de creare şi publicare a informaţiei nu este neapărat remunerată – ba chiar, tipic vorbind, rămâne neplătită.

După cum observă Adrian Daub în analiza din cartea sa *What Tech Calls Thinking: An Inquiry into the Intellectual Bedrock of Silicon Valley*[25], ideologia Silicon Valley este bazată pe devalorizarea conţinutului în favoarea platformei. Profitul, în această accepţiune, trebuie să meargă la cei care construiesc tehnologiile ce fac posibilă *circulaţia* mesajelor, şi nu la cei care *generează* conţinutul fără de care aceste platforme nu ar avea sens. În ciuda acestei devalorizări constante a contribuţiilor individuale şi a

extragerii de valoare financiară în favoarea platformelor, numărul de utilizatori și contribuitori a crescut permanent, de la an la an.

Pe măsură ce diversitatea de informații și voci s-a transformat în abundență, *filtrele algoritmice* au ajuns să aibă o funcție vitală în spațiul digital. Nici un om nu poate cuprinde totalitatea informațională a internetului, nici măcar informațiile generate la nivelul unui minut sau al unei secunde (vezi figurile 29 și 30). Aceste filtre constau în motoare de căutare, algoritmi care ne recomandă noi piste de vizionare sau selectează ce mesaje sau știri să primim în *feed*-ul rețelelor de socializare pe care activăm. În măsura în care aceste filtre sunt create și optimizate de corporații care urmăresc expansiunea continuă a bazei lor de utilizatori și creșterea accelerată a profitului, apar conflicte inevitabile între logica adevărului și logica persuasiunii.

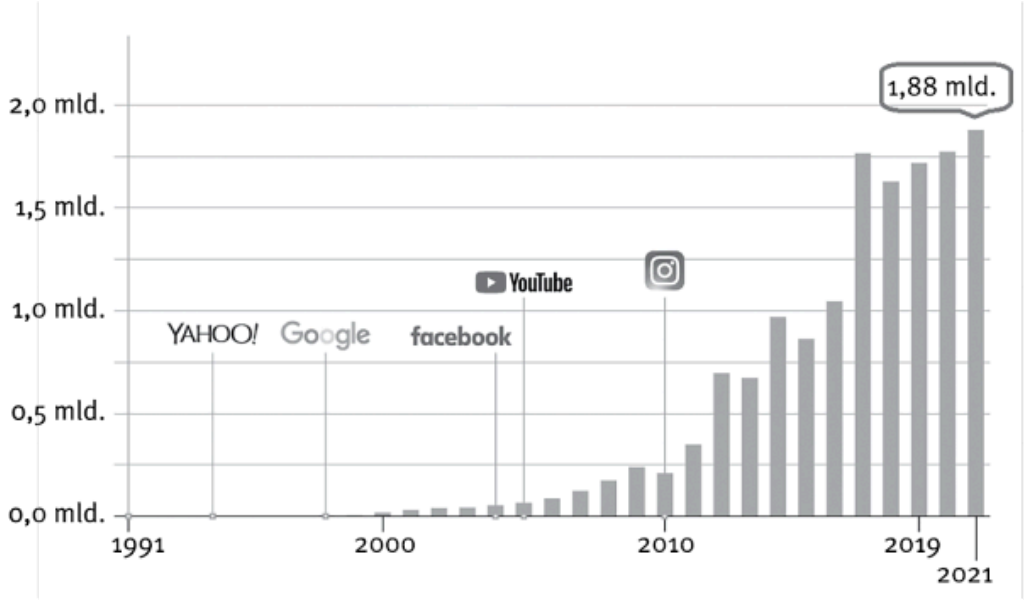


Fig. 29 Câte site-uri există? Evoluția internetului măsurată prin numărul total de site-uri, 1991–2021. Sursa: Statista.

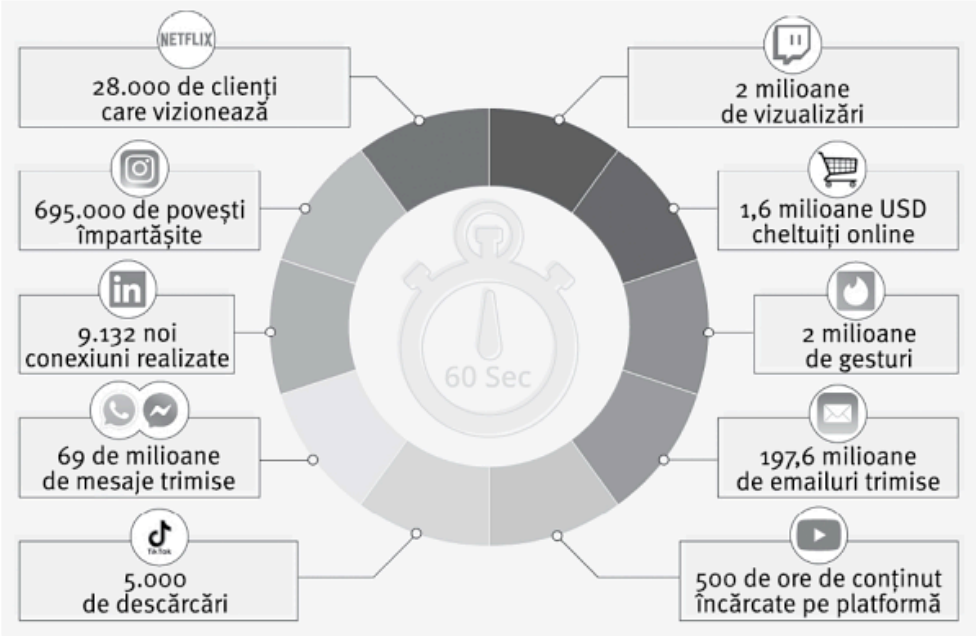

NETFLIX
28.000 de clienți care vizionează
2 milioane de vizualizări
695.000 de povești împărtășite
1,6 milioane USD cheltuiți online
9.132 noi conexiuni realizate
2 milioane de gesturi
60 Sec
69 de milioane de mesaje trimise
197,6 milioane de emailuri trimise
5.000 de descărcări
500 de ore de conținut încărcate pe platformă

Fig. 30 Volumul estimat de date create online într-un minut în 2021. Sursa: Statista.

O altă zonă de expansiune a mediului online și de creștere a informațiilor produse și procesate acolo este internetul lucrurilor (*Internet of Things* sau *IoT*). De la an la an, tot mai multe obiecte devin *smart*, adică incorporează senzori și procesoare, colectând datele utilizatorilor și transmițându-le mai departe, spre producătorii lor și adesea spre brokeri de date. Vorbim aici atât de obiecte gospodărești, precum frigidere sau sisteme de iluminat casa, care pot fi integrate în sisteme *smart-home*, cât și de obiecte de uz medical, militar sau industrial.

Dacă în 2019 erau aproximativ 7,7 miliarde de dispozitive conectate în internetul lucrurilor, estimările spun că volumul acestora se va tripla până în 2030.[26] Toate aceste obiecte colectează informații din lumea fizică, constituind o vastă rețea de monitorizare și modelare a realității, inclusiv a vieții noastre. De asemenea, toate aceste obiecte participă la un sistem tot mai autonom de decizii automatizate. Provocările sunt enorme, mai ales privind asigurarea securității. Sistemele de operare și aplicațiile utilizate de aceste obiecte sunt diverse, adesea puțin compatibile unele cu altele, testate în pripă și neprotejate de soluțiile standard ale internetului computerelor, precum programele antivirus.

Potențialul catastrofic al unui superinternet în care circuite

informaționale intangibile sunt legate de structuri materiale și corpuri biologice, într-un flux continuu și direct, este sesizat de Bruce Schneier în cartea sa *Click Here to Kill Everybody: Security and Survival in a Hyper-Connected World*.[27] Într-un astfel de sistem tot mai complex, informațiile culese și transmise sunt vulnerabile la erori și atacuri tot mai diverse. Până acum am subliniat riscurile aduse de dominația logicii capitaliste în filtrarea cunoașterii digitale. În schimb, internetul lucrurilor aduce în prim-plan vulnerabilitatea corpurilor și infrastructurilor noastre la atacuri malițioase, dar și la erori aleatorii propagate în sistem.

Complexitatea soluțiilor de inteligență artificială conduce la paradoxuri. Intervenții minimaliste, dar bine gândite, pot afecta semnificativ deciziile sistemelor de IA. Astfel de atacuri sunt, așadar, accesibile și hackerilor cu puține resurse materiale, dar care au o oarecare ingeniozitate sau perseverență. De exemplu, în așa-numitele *one pixel attack*, modificarea intenționată a unui singur pixel poate da peste cap clasificarea realizată de un model de inteligență artificială. Consecințele pot fi catastrofale pentru deciziile luate de dispozitivul pe care-l conduce.[28] Dacă, așa cum arată autorii, o inteligență artificială poate vedea un cal ca pe o broască sau un cerb ca pe un avion, riscul unei manipulări cu consecințe grave nu este greu de înțeles.

Bunăoară, mașinile Tesla au fost ținta unor știri de senzație privind impactul atacurilor asupra deciziilor automate. În aprilie 2019, cercetătorii din laboratorul Keen Security, al companiei Tencent, anunțau că pot convinge o Tesla să intre pe contrasens cu ajutorul câtorva abțibilduri strategic plasate pe carosabil.[29] Experimentul lor se desfășurase în condiții controlate, nu pe șosele din lumea reală. Tesla a recunoscut vulnerabilitatea, (re)amintind publicului larg că pentru așa-numitul autopilot este nevoie în continuare de un șofer care să fie atent la trafic și să poată prelua oricând controlul.

Există, de altfel, o discrepanță continuă între marketingul Tesla, care glorifică autonomia mașinii, și manualul de utilizare, care solicită supervizarea umană pentru a evita posibilele accidente și pentru a prelua responsabilitatea. Mașinile Tesla au fost implicate în accidente fatale provocate de pilotul automat. Așadar, omul este încă un ingredient esențial,

cel puțin pentru a prelua vina.[30]

În noiembrie 2020, publicația *Wired* raporta posibilitatea existenței unor semne de circulație fantomă: atacatorii pot proiecta imagini efemere pe asfalt sau pe afișaje electronice comerciale, convingând mașina să se oprească brusc.[31] Afișajele electronice de pe autostrăzi sunt relativ ușor de modificat chiar și de la distanță, putând fi folosite, ipotetic, pentru a dezorganiza grav traficul. De exemplu, pe lângă faptul că ele pot opri mașinile, la fel de important este că nu lasă nici o dovadă a atacului.

Astfel de atacuri simple pot afecta și decizii militare sau medicale.[32] Prin urmare, abundența fluxurilor informaționale ridică nu doar provocări legate de selecția cunoașterii oferite oamenilor, ci și riscuri privind manipularea cunoașterii oferite obiectelor.

Cunoașterea autonomă și pseudoumană: inteligența artificială

Unde nu-i cap, vai de picioare.

Tehnologia informației ilustrează foarte bine trecerea de la o acumulare cantitativă la o transformare calitativă, formulată încă de acum două sute de ani de Hegel și invocată apoi de Marx, Engels și alți observatori ai transformărilor sociale radicale.[33] În ultimii ani, modelele de inteligență artificială care prelucrează vocea, imaginile sau textul au făcut progrese greu de imaginat în trecut. Modelele bazate pe învățare automată pot acum citi, traduce, sintetiza și chiar scrie text, pot recunoaște fețe umane, pot crea muzică, pot juca roluri de actori în reclame, pot clasifica artefacte arheologice și pot interpreta radiografii, printre altele – cu rezultate comparabile calitativ și covârșitoare cantitativ față de cele obținute de oameni.[34]

Prin urmare, suntem martorii unei extinderi continue a utilizării clasificărilor și deciziilor algoritmice în domenii diverse: de la marketingul predictiv și algoritmii de recomandări în realizarea de asigurări la recrutarea și monitorizarea forței de muncă, sănătate, educație, ordine publică etc. Cuvântul *smart* a ajuns să descrie mai degrabă un obiect digital decât o persoană, tot așa cum *smart government* sau *smart medicine* fac mai degrabă trimitere la încorporarea inteligenței artificiale decât la ascuțirea

celei a experţilor umani.

Delegarea capacităţilor de percepţie, clasificare, diagnoză şi decizie către sisteme automate ridică semne de întrebare cu privire la asumarea responsabilităţii finale pentru aceste decizii.[35] Cine poate fi tras la răspundere retroactiv şi proactiv pentru deciziile luate de sisteme automate cu impact direct asupra oamenilor? Mai mult decât atât, vom asista şi la *o delegare a competenţelor noastre morale* către aceste reţele de difuziune a responsabilităţii între oameni, organizaţii şi maşini?[36]

„Dacă vrei să îngropi o problemă, fă o comisie“, spune un proverb contemporan. Deciziile colective reprezintă o soluţie bine-cunoscută de diluare a efortului moral personal şi de evitare a responsabilităţii. Dacă pentru orice hotărâre dificilă este mai simplu să creăm o comisie, vom ajunge oare să scăpăm de povara celor mai sensibile probleme dând vina pe boţi ce nu pot fi traşi la răspundere?

Cunoaşterea-combustibil: de la informaţie, la acceleraţie

Unde dai şi unde crapă.

În 1973, într-o declaraţie în faţa Congresului SUA, economistul american Kenneth Boulding remarca: „Oricine crede că o creştere exponenţială este posibilă pe termen infinit într-o lume finită este fie nebun, fie economist“[37]. Jumătate de secol mai târziu, această perspectivă descrie în continuare dinamica economiei globale.

Între timp, au survenit creşteri semnificative în eficienţă care au permis o decuplare relativă a creşterii economice de sporirea consumului de resurse, mai ales în economii dezvoltate precum cele din Uniunea Europeană şi SUA.[38] Totuşi, la nivel global, consumul de resurse şi emisiile de carbon continuă să urce (vezi figura 31). Tehnologia informaţiei reprezintă o parte a acestui tipar.

În economia ultimelor decenii, asistăm la o buclă de feedback pozitiv, adică de intensificare reciprocă, între dezvoltarea universului imaterial al biţilor şi consumul resurselor terestre, necesare pentru a produce, a folosi şi a arunca la gunoi dispozitivele digitale.[39] Noile tehnologii ale cunoaşterii, mai ales modelele de învăţare automată, sunt tot mai mari consumatoare de

resurse financiare și energetice.[40] Nu doar consumul lor propriu-zis contribuie la această accelerare, ci și toate celelalte industrii care sunt stimulate de progresul tehnologiilor digitale, producând mai mult pentru mai mulți consumatori.

Pe de altă parte, omenirea își pune speranțe în potențialul IA de eficientizare a economiilor și de combatere a încălzirii globale. Modelele de învățare automată pot îmbunătăți predicția climatică, inclusiv a evenimentelor extreme. Ele pot susține dezvoltarea unor soluții noi de producere a energiei, de îndepărtare a carbonului din atmosferă, de monitorizare și de eficientizare a consumului de energie – de la clădiri, la lanțuri de distribuție sau rute de transport.[41] Întrebarea rămâne: în ce măsură vom reuși să folosim preaplinul cunoașterii digitale pentru a compensa distrugerea accelerată a planetei? Sau, cu alte cuvinte, cum vom putea să prevenim transformarea abundenței într-o combinație devastatoare de exces și deficit?

În paginile care urmează, vom explora în mai mare detaliu această ambivalență a cunoașterii generate și intermediate digital, precum și impactul său asupra vieții umane și dreptății sociale. Fiecare tehnologie intrată pe piață ajunge să fie un instrument în mâna a mii, milioane sau miliarde de utilizatori – indivizi, organizații, state sau alte tipuri de colectivități ce-și urmăresc feluritele interese. Adesea, creatorii tehnologiei o rafinează și îi adaptează distribuția pentru a-și maximiza profitul. Utilizatorii, la rândul lor, încearcă să-și extindă controlul asupra lumii, folosind-o în moduri prevăzute și neprevăzute. Produsele tehnologice au anumite disponibilități (*affordances*), adică stimulează sau încurajează anumite acțiuni atât prin destinația pe care o au, cât și prin diferite proprietăți pe care le descoperă cei ce le folosesc, experimentând cu ele.

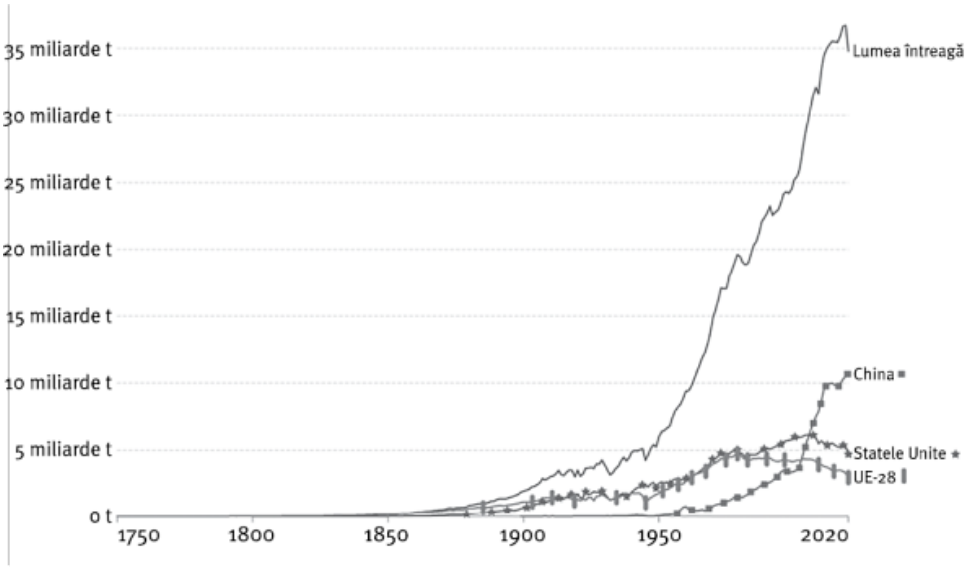


Fig. 31 Evoluția emisiilor anuale de CO2 în diferite regiuni ale globului, în perioada 1750–2020. Sursa: Our World in Data.

Într-o analiză a disponibilităţilor digitalizării în educaţie, Grainne Conole şi Martin Dyke[42] identifică mai multe direcţii spre care ne împing tehnologiile. Printre acestea se numără *accesibilitatea* tot mai crescută a unor cantităţi vaste de informaţie, o *diversitate* uimitoare de experienţe, putința de *comunicare și colaborare* dincolo de granițele de timp și spațiu ale lumii fizice, posibilitatea de a stimula *reflecția* (dar și de a o împiedica) prin *viteza* tot mai mare a interacţiunilor și *stimularea continuă*. Noul univers informaţional este *multi-modal și non-liniar*, purtându-ne în zigzag de la text la imagine, de la audio la video sau la suprapuneri de media. Conole și Dyke remarcă însă și înlesnirea unor noi relații de putere și control prin *monopolurile* pe care marile firme de tehnologie le consolidează și prin *supravegherea* continuă la care deja ne-am obișnuit să fim expuși. Anumite disponibilităţi care par să aducă numai lucruri bune sfârșesc prin a avea și efecte negative neanticipate. De exemplu, comunicarea instantanee gratuită la nivel mondial, oferită de e-mail și de mesageriile digitale, ajunge să schimbe normele și sancţiunile sociale privind temporalitatea conversaţiilor. Resimțim adesea, cu frustrare sau vinovăție, presiunea de a răspunde rapid, ba chiar imediat la mesajele amicilor, colegilor sau șefilor. Toleranța pentru întârzierea răspunsurilor digitale scade, în același timp în care abundența comunicării ne copleșește

cu fluxuri interminabile de scrisori și replici. Abundența poate deveni intolerabilă atunci când se transformă în exces.

---

6. William Gibson, citat în *The Economist*, 4 decembrie 2003.

7. Disponibil online la https://delphi.allenai.org.

8. Vezi estimările pe site-ul SeedScientific: https://seedscientific. com/how-much-data-is-created-every-day.

9. Data căutării: 17 octombrie 2021.

3. Excesul cunoașterii în societatea algoritmică

Unde e un car cu înțelepciune, acolo sunt două de nebunie.

Faptul că erorile și limitările cunoașterii pot provoca daune morale este aproape de la sine înțeles și e adesea întâlnit în reflecția publică. În capitolul următor, voi discuta despre implicațiile etice ale deficitului de cunoaștere a soluțiilor bazate pe Big Data.

În ce măsură însă *excesul* de cunoaștere poate fi problematic din punct de vedere etic? În cazul deciziilor algoritmice și Big Data, unele dintre criticile importante aduse acestor evoluții tehnologice și organizaționale pornesc nu de la lipsa validității cunoașterii ci, dimpotrivă, de la un anumit surplus a acestei cunoașteri. Acest surplus are două dimensiuni distincte. Pe de o parte, este vorba de exces în formă *absolută*: cantități sau calități ale informației care apar atunci când nu ar trebui să apară și degradează calitatea judecăților morale. Pe de altă parte, excesul poate fi *relativ*, sub forma asimetriilor de informație între parteneri în interacțiune. Când un agent știe prea mult, iar altul prea puțin, calitatea interacțiunii poate fi viciată.

Tony Smith este unul dintre milioanele de americani care suferă de sindromul de apnee în somn.[1] În timpul somnului, respirația sa se oprește frecvent, ceea ce îl forțează să se trezească de zeci de ori pe noapte și să fie somnolent toată ziua. Desigur, apneea sa duce la probleme la serviciu și la riscuri de accident în trafic, punându-i traiul și chiar viața în pericol.

Există însă o soluție pe care Tony Smith a adoptat-o cu entuziasm: un aparat special care egalizează presiunea aerului la inspirație și expirație numit CPAP (*Continuous Positive Airway Pressure*). Acest dispozitiv transmite însă cu succes nu doar aer presurizat, ci și date – adesea, fără chiar ca utilizatorii să fie conștienți. În cazul lui Tony Smith, modelul mai vechi pe care îl avusese stoca datele pe un card pe care îl putea duce la doctor pentru a examina calitatea somnului și eficiența aparatului. Datele sale erau comunicate doar medicului. Versiunea mai nouă și *smart* pe care a

obținut-o, conectată la WiFi, transmitea însă datele sale nu doar medicului, ci și companiei producătoare. Smith a fost șocat când și-a dat seama de acest lucru, după ce a primit un mail de la firma respectivă prin care îl felicita pentru prima noapte de terapie cu noul aparat. De unde știa producătorul aparatului că Smith îl folosise? Nu a durat mult până când și-a dat seama că toate informațiile privind somnul său erau înregistrate și transmise producătorului, firmei care produce consumabilele pentru aparat și chiar companiei de asigurări medicale. Aceste date pot fi și sunt adesea folosite împotriva intereselor pacientului, arată ProPublica în aceeași analiză. Companiile de asigurări invocă informațiile obținute pentru a refuza pacienților decontarea costurilor aparatelor, în conflict cu recomandările medicilor curanți. Interesele companiilor de asigurări sunt, desigur, diferite față de interesele pacienților sau cele ale doctorilor, iar circulația datelor joacă un rol foarte important în echilibrul de forțe dintre acești actori. Conform legislației din SUA, doctorii și spitalele nu pot transmite datele medicale ale pacienților unor terțe părți, pentru a-i feri de abuzuri. Totuși, legislația nu acoperă și informațiile colectate de firmele producătoare de dispozitive medicale, care nu sunt furnizoare de servicii de sănătate, și intră deci în altă categorie – chiar dacă, în practică, ele adună tot date medicale.

Prin urmare, producătorii aparatelor *smart* conectate la internetul lucrurilor colectează datele și le vând unor terțe părți, inclusiv companiilor de asigurări.[2] Adesea, consumatorii semnează rapid că sunt de acord cu condițiile de utilizare și diseminare a datelor lor fără să înțeleagă cu adevărat ce informații sunt colectate despre ei și cui vor fi acestea transmise. Acest flux de informații între parteneri asimetrici, anume pacientul, pe de o parte, și compania de asigurări, pe de altă parte, nu se referă doar la mașinăriile CPAP. Multe alte dispozitive pot culege date foarte sensibile privind starea de sănătate și stilul de viață al unei persoane, date care să fie apoi folosite, de exemplu, pentru a-i stabili eligibilitatea sau prețul pentru asigurări de sănătate. Printre multiplele dispozitive medicale inteligente care-și pot spiona utilizatorii în favoarea asiguratorilor se află și monitoarele cardiace, glucometrele și brățările care monitorizează activitatea fizică, somnul și caloriile.[3]

Mihai Stelian Rusu, sociolog la Universitatea Lucian Blaga din Sibiu, caracterizează bârfa ca o conversaţie informală despre oameni concreţi, absenţi, cu conţinut evaluativ şi adesea negativ, care foloseşte intereselor participanţilor la schimbul de informaţii, făcând abstracţie de interesele ţintei sau chiar fiind dăunătoare acestora.[4] Putem spune oare că dispozitivele *smart* pe care le folosim pe post de asistenţi sau consilieri digitali ne bârfesc cu producătorii lor şi cu brokerii de date, cărora le şoptesc vrute şi nevrute despre acţiunile noastre? Poate că aceasta este o antropomorfizare excesivă a tehnologiei. Totuşi, cele două scenarii au în comun asimetria de acces şi participare la conversaţie, care îi lasă pe utilizatori descoperiţi în faţa unui flux de cunoaştere *despre* ei care se poate întoarce, pe nepusă masă, *împotriva* lor.

În cele ce urmează voi discuta mai multe *faţete ale excesului de cunoaştere*:

– *excesul de precizie*, manifestat prin accesul organizaţiilor comerciale, politice sau statale la informaţii detaliate despre indivizi (clienţi, votanţi, cetăţeni sau alţi indivizi de interes), informaţii care permit microţintirea unor categorii nişate de public prin măsuri comerciale, de factură politică sau de supraveghere statală;

– *excesul de accesibilitate a informaţiilor*, manifestat prin disponibilitatea persistentă pe internet a detaliilor din viaţa noastră digitală, chiar şi după ce acestea au devenit irelevante;

– *excesul de informaţii ilegitime*, precum genul, vârsta sau rasa/etnia unei persoane, care pot conduce la discriminare în decizii algoritmice;

– *excesul costurilor* financiare, ecologice şi psihologice ale soluţiilor Big Data;

– nu în ultimul rând, voi discuta şi *excesul relativ*, adică dezechilibrul cunoaşterii, specific capitalismului datelor, care are şi implicaţii etice.

Excesul de precizie: dezanonimizarea datelor şi daunele voluntare sau involuntare asupra indivizilor

Nu te lăuda celui ce te ştie.

Ni se întâmplă uneori să primim online reclame extraordinar de precise,

pentru produse la care ne-am gândit sau despre care tocmai am vorbit cu cineva. Dacă nu ni s-a întâmplat, putem găsi pe internet numeroase anecdote despre astfel de cazuri. Aceste coincidențe, care nu par deloc astfel, ne izbesc ca fiind stranii și ne dau fiori. Să trecem în revistă câteva exemple. Oscar a văzut pe Instagram o reclamă la un cuțit japonez Masomoto KS, exact după ce primise cadou un cuţit identic de la un prieten, fără să fi căutat vreodată ceva similar pe internet.[5] Dan a fost uimit de o reclamă la un produs de nișă – un topitor de brânză Raclette –, pe care s-a dovedit că şi-l dorește, chiar dacă, la acel moment, (încă) nu știa la ce folosește.[6] Carl a văzut pe Facebook o reclamă pentru un obiectiv de fotografiat Rokinon imediat după ce îi fusese menționat de chiropracticianul său, fără ca el însuși să fi căutat așa ceva înainte.[7] Neville a început să primească reclame la mâncare de pisici, chiar dacă nu are pisici și nu a căutat așa ceva online, dar subiectul apăruse într-o discuție cu soția, iar Isaac s-a pomenit cu reclame la fierbătoare de ouă după ce începuse să se gândească la o posibilă schimbare de dietă.[8] Artiștii Tega Brain și Sam Lavigne au colectat sute de astfel de istorioare în proiectul lor „Noile Organe“, prin care explorează urechile digitale ce ne urmăresc fiecare gest online.[9]

Unii oameni cred că Facebook, Amazon și alte companii ne spionează efectiv prin microfonul telefonului, activându-l fără știrea noastră și înregistrând ce spunem. Deși această idee mai circulă încă, multiple investigații indică faptul că este falsă.[10] Spionajul audio ar avea costuri foarte mari pentru stocarea și analiza datelor. Totodată, meritele ar fi foarte mici din cauza dificultății identificării și separării informațiilor relevante de zgomotul ambiental și de alte distorsiuni, precum ironia, gluma, dezacordul sau referințele la opinii ale altor persoane. Mai mult, bogăția datelor relevante disponibile prin urmărirea activității noastre online și offline arată că o astfel de supraveghere audio nici nu este necesară pentru stabilirea profilului consumatorilor.

Prin urmare, deși pare că suntem ascultați, actualmente nu mai este nevoie de microfoane secrete pentru a anticipa cu precizie uimitoare preferințe, intenții și gânduri. La urma urmei, fiecare click pe care îl facem online, fiecare secundă pe care o petrecem pe vreun site, fiecare

cumpărătură pe internet, dar și fizic (cu cardul de credit sau cu un card de fidelitate), sunt înregistrate și acumulate într-un gigantic profil digital personal pe care Big Tech îl are la îndemână. Nu este de mirare că dorințele noastre nu mai sunt un mister, decât poate pentru noi înșine.

Nu lipsesc nici vocile, mult mai rare, care consideră că toată această superprecizie a reclamelor personalizate este, de fapt, o iluzie întreținută de Big Tech pentru a-și putea vinde produsele. Cu alte cuvinte, ceea ce pare exces ar fi, de fapt, un deficit. Unele studii privind acuratețea datelor acumulate de brokeri arată erori frecvente de clasificare și o variabilitate ridicată a preciziei, respectiv impreciziei[11]. Augustine Fou observă că, adesea, primim reclame insistente la produse pe care tocmai le-am cumpărat – ceea ce nu denotă o buclă de feedback prea bună.[12] În plus, observă el, tot mai mulți oameni folosesc programe care blochează reclamele online (*ad-blockers*), scăpând din circuitul AdTech.

Simultan, există o puzderie de boți care imită oameni și se uită sau fac *click* pe reclame, pentru a da companiilor senzația că marketingul personalizat este eficient și pentru a le lua banii. Astfel de boți ajung, inevitabil, să polueze bazele de date în care ei se pretind consumatori legitimi. În afară de boți, există și alte tehnici de fraudare, unele mizând chiar pe implicarea participanților umani, cum este cazul fermelor de clickuri, sau pe combinații uman-tech, precum hackeri ce folosesc diferite aplicații. Toate aceste tehnici de fraudare duc la pierderi de miliarde de dolari anual în industria AdTech[13], dar și în alte industrii care depind de click-uri ușor de falsificat. Bunăoară, completarea de chestionare online[14] sau chiar recrutarea de candidați pentru universități în căutare de studenți[15] au devenit activități care nu mai reflectă doar inputuri umane.

Chiar dacă acceptăm că boții și alte surse de erori au un rol semnificativ în capitalismul datelor, potențialul companiilor de a ne identifica, monitoriza și prezice acțiunile pe baza profilurilor noastre digitale rămâne îngrijorător. Vom explora în continuare latura excesivă a acestui potențial, urmând să discutăm reversul medaliei, adică deficitul, în capitolul următor.

Unul dintre riscurile morale asociate cu predicțiile bazate pe Big Data se referă chiar la granularitatea cunoașterii despre indivizi pe care aceste soluții o oferă firmelor sau statelor. Oamenii sunt, într-o oarecare măsură,

predictibili pentru cei care îi cunosc. Cei apropiați, fie că este vorba despre prieteni sau despre familie, vecini, colegi sau alte persoane din relații de proximitate socială și spațială, pot ghici destul de bine preferințele noastre și chiar acțiuni pe care le vom face. Pentru persoanele cu care interacționăm mai mult sau mai puțin des, dar relativ constant, astfel de trăsături personale (precum cu cine votăm, care sunt credințele noastre religioase, care sunt preferințele noastre sexuale, ce ne place să mâncam sau cu ce branduri ne identificăm) sunt destul de clare.

Totuși, până nu demult, aceste informații erau, de regulă, limitate la un cerc destul de redus de oameni. Situațiile în care străinii ar fi știut astfel de lucruri despre un individ obișnuit erau rare – presupunând activități de spionaj sau investigații jurnalistice sau alte tipuri de culegere de informații rezervate persoanelor cu un profil public sau profesional deosebit. *În era Big Data, trăim un paradox: a devenit posibil ca un profil personal și detaliat al nostru să fie accesibil unor străini prin intermediul datelor digitale stocate despre noi, profil care însă nu ne este accesibil chiar nouă, utilizatorilor individuali, cei care generăm datele*. Intimitatea datelor personale este încălcată pe măsură ce volume tot mai mari de detalii despre cuvintele-cheie ale căutărilor noastre, cumpărăturile noastre, clickurile și like-urile noastre sunt agregate într-un nor de informații care ne descrie preferințele și evoluția personală.

Acest nor de informații permite formularea unor predicții privind acțiunile noastre, predicții bazate pe cunoașterea unor vulnerabilități sau dispoziții personale de care putem să nu fim conștienți. De aici derivă anxietatea prezentă în anumite dezbateri publice privind capitalismul datelor, unde se discută felul în care în viitor agenții care vor controla datele noastre personale ne vor cunoaște mai bine decât ne cunoaștem noi înșine.

Oamenii sunt înclinați să se să se vadă într-o lumină favorabilă, făcând erori sistematice în înțelegerea pe care o au asupra propriei persoane și, mai ales, asupra propriilor vulnerabilități. Este oare posibil ca organizații comerciale să cunoască mai bine decât noi înșine momentul în care cedăm unor tentații precum consumul de alcool, tutun ori dulciuri sau altor comportamente riscante sau vulnerabilități personale despre care noi credem, prea optimist, că le putem controla? Este oare posibil ca această

diagnoză a limitărilor personale să fie folosită împotriva intereselor noastre? Răspunsul pare să fie afirmativ în ambele privințe.

Mai mult, uneori oamenii au secrete pe care nu le împărtășesc nici celor apropiați. Aceste secrete pot include informații considerate periculoase sau surse de vulnerabilități și riscuri de stigmatizare. De exemplu, detaliile privind orientarea sexuală, boli fizice sau psihologice, orientări politice minoritare sau interzise, credințe religioase atipice sau condamnate etc. pot conduce la condamnarea, formală sau informală, a unei persoane. Chiar și cei mai apropiați oameni pot să nu cunoască astfel de lucruri despre noi, dacă le ținem secrete. Totuși, aceste taine pot fi detectate dacă un observator are acces la volume vaste de date despre acțiunile noastre, inclusiv la comportamente considerate private, precum căutarea de informații pe net, din intimitatea propriei locuințe, corespondența electronică sau citirea și vizionarea anumitor materiale. Astfel, ceea ce era anterior relativ ușor de păstrat secret, acum riscă să devină un secret al lui Polichinelle.

Acest exces de vizibilitate ridică probleme privind autonomia decizională a indivizilor, care pot fi manipulați sau șantajați de organizațiile ce dispun de astfel de informații. Fie că vorbim despre microțintirea votanților prin reclame electorale, microțintirea consumatorilor prin publicitate sau monitorizarea posibililor oponenți politici, ne situăm pe un spectru larg al restrângerii sau chiar al anulării libertății individuale de decizie în favoarea intereselor părților ce controlează informația. Pe măsură ce tehnologiile folosite în mediul medical sau în viața cotidiană se bazează tot mai mult pe date[16], profilurile noastre vor deveni tot mai puțin intime și mai pasibile de a ne vulnerabiliza.

Teoretizarea capitalismului supravegherii[17] surprinde tocmai această transformare a surselor valorii economice. Zuboff pornește de la evoluția modelului de afaceri al lui Google, identificând emergența unui nou tip de resursă: *surplusul comportamental*, disponibil celor ce controlează volume tot mai vaste de date privind acțiunile noastre.

Pe baza acestei noi materii prime, giganți tehnologici, precum Google, Facebook sau Amazon, dar și alte companii care cumpără și procesează fluxuri de date privind acțiunile oamenilor, elaborează produse de predicție. Aceste produse pot avea în vedere anticipările deciziilor noastre ce țin de

consum, de vot sau de interacţiune socială pe platforme de socializare, în cadrul aplicaţiilor de *dating* etc. Apare şi un proces specific de acumulare bazat pe această nouă formă de capital, o acumulare care conduce la consolidarea dezechilibrelor de putere. După cum subliniază Zuboff, problema cu aceste produse de predicţie este că nu încearcă doar să anticipeze acţiunile noastre ci mai mult să le influenţeze şi să le ducă în direcţia dorită de companiile ce controlează datele. Desigur, dintotdeauna firmele, partidele politice şi alte organizaţii au dorit să anticipeze şi să influenţeze comportamentul consumatorilor, electoratului etc. Problema cu noua organizare bazată pe Big Data constă în asimetria informaţională tot mai crescută între cei care au acces la date şi cei care sunt descrişi şi modelaţi prin date, aşa cum voi discuta mai jos.

Dimensiuni ale excesului de precizie

Buturuga mică răstoarnă carul mare.

Pentru a analiza mai detaliat implicaţiile etice ale excesului de precizie, voi evidenţia o serie de dimensiuni care caracterizează acest tip de exces de cunoaştere: dimensiunea temporală, cea experimentală, cea corelaţională şi cea identitară. Aceste dimensiuni nu sunt exhaustive, dar permit scoaterea în evidenţă atât a avantajelor, cât şi a riscurilor la care se expun companiile Big Tech, precum şi impactul lor asupra indivizilor, care sunt supuşi, vrând-nevrând, la conturarea identităţilor lor digitale.

Excesul de precizie are o dimensiune temporală. Aceste predicţii pot fi livrate la momentul oportun din zi sau din noapte prin intermediul platformelor digitale personalizate, exploatând o serie de vulnerabilităţi personale la care, în trecut, firmele nu aveau acces. Acest exces de precizie şi de intervenţie poate fi găsit în conceptualizarea micromomentelor[18]. Prin amploarea şi granularitatea fină a datelor disponibile despre clienţi, Google promite companiilor o inserare a anunţurilor publicitare în micromomentul decizional în care indivizii sunt cel mai înclinaţi să treacă la acţiune şi să se orienteze spre ofertantul pe care îl văd. Acest tip de precizie temporală este posibilă doar în condiţiile omniprezenţei mediului digital, care a devenit parte dintr-o realitate fizică augmentată ce permite cunoaşterea şi

influențarea comportamentelor online, dar și offline, prin stimulări digitale.

Excesul de precizie are o dimensiune experimentală. Fie că vrem sau nu, fie că suntem conștienți sau nu, în orice moment de navigare online putem fi participanți într-un experiment. Așa-numita testare A/B sau experimentare segmentată (*A/B testing* sau *split testing*) se referă la practica de îmbunătățire continuă a designului unei platforme prin testarea fiecărei mici trăsături pe două sau mai multe grupe de utilizatori. Prin astfel de comparații, se poate vedea care versiune stârnește rezultatele cele mai favorabile companiei. Acestea nu sunt experimente științifice și echipele de proiectare digitală nu cer consimțământul participanților. Acest așa-zis acord de participare este, oricum, oferit implicit, undeva prin „Termenii și condițiile de utilizare“ ai serviciului, pe care îi acceptăm de regulă fără să-i citim. Testarea A/B poate fi folosită atât în cazul oricărui tip de decizie: pentru a alege culorile care țin mai mult timp utilizatorii pe un site, pentru a desemna cea mai persuasivă versiune a unei reclame electorale pentru un anumit segment de public[19] sau pentru a hotărî forma unui buton care va atrage cele mai multe clickuri[20].

Marile companii de tech au acces la bugete enorme pentru a testa ce design este cel mai persuasiv, câștigând astfel un avantaj în raport cu utilizatorii, care nu pot obține o astfel de cunoaștere. Facebook face experimente masive pe noi pentru a vedea dacă ne poate influența emoțiile sau cum ne poate face să cedăm mai multe date personale.[21] Noi nu facem experimente pe Facebook – sau, dacă facem, sunt la scară redusă, testând algoritmii de aici și acum. Așadar, este cel mai probabil ca, în acest caz, puterea de modificare a comportamentelor să se amplifice unidirecțional.

Excesul de precizie are și o dimensiune corelațională. De regulă, putem prezice în linii mari comportamentul consumatorilor sau votanților în funcție de trăsăturile lor socio-demografice sau în funcție de alte comportamente disponibile. Big Data face posibilă identificarea unor predictori, sau *proxy*, mult mai distanți și mai ne-intuitivi, precum și combinarea lor pentru a decupa audiențe receptive pentru un anumit mesaj. Desigur, putem vorbi aici despre un deficit de înțelegere, deoarece acești factori predictivi nu sunt cauze ale fenomenului de interes. Totuși, chiar și așa, pot oferi o pârghie de acțiune prin clasificarea tipurilor de votanți,

clienți etc., permițând o segmentare fină și dinamică a audienței.

Este important să subliniem schimbarea calitativă care apare din creșterea cantitativă indusă de Big Data. În cazul modelelor statistice clasice, apelul la corelații iluzorii, care nu erau susținute de relații cauzale, era considerat dăunător, deoarece putea conduce la predicții eronate. În era datelor vaste, pot fi găsite atât de multe corelații între atât de multe variabile, încât apare un tip de descriere statistică densă. În acest caz, multiplicitatea relațiilor de coapariție a anumitor gesturi asigură, pe termen scurt, validitatea predictivă a modelelor (sau, cel puțin, acesta este discursul care justifică știința datelor). Corelații aparent triviale, de exemplu, între preferința pentru un anumit tip de înghețată și comportamentul electoral, ar fi fost neutilizabile, ca atare, în analizele statistice clasice – dar acestea devin materie primă valoroasă pentru modelele de analiză a datelor. Mai mult, aceste analize nu sunt orientate doar spre trecut, ci și spre țintirea unor acțiuni de influențare a opiniei publice, precum și a faptelor propriu-zise ale publicului. Indivizilor le pot fi oferite mesaje personalizate sau stimulente personalizate croite în funcție de preferințele lor recente și detaliate, sporind puterea de manipulare a celor care au acces la astfel de date.[22]

Unul dintre episoadele reprezentative pentru riscul asociat cu dimensiunea corelațională a excesului de precizie îl constituie campania lanțului de magazine Target, desfășurată în SUA în 2012. Acest episod este prezentat într-un articol din *The New York Times* dedicat predicțiilor în marketing[23], și se referă la cazul unui tată care a venit supărat la magazin deoarece Target îi trimitea fiicei lui adolescente oferte de reducere pentru produse destinate tinerelor mame. Câteva zile mai târziu, tatăl a aflat că fiica lui era, de fapt, gravidă. Această întorsătură de situație, în care un supermarket ajunge să cunoască detalii intime despre o persoană înaintea celor apropriați, a devenit simptomul excesului de vizibilitate pe care fluxurile de date personale îl oferă organizațiilor despre viața noastră, precum și al riscurilor pe care această vizibilitate le implică.

Trăsăturile noastre psihologice, precum și preferințele sau dispozițiile comportamentale, inclusiv orientările sexuale, politice, religioase etc., pot fi inferate pe baza comportamentelor noastre publice online. Pornind de la like-urile pe care le dă o persoană pe rețelele de socializare și asocierea lor

anterioară cu trăsături de personalitate, bazate pe studii corelaţionale ample, observatorii care dispun de date îi pot detecta caracteristicile personale, precum genul, etnia, religiozitatea, orientarea sexuală, preferinţele politice, trăsăturile de personalitate, dar şi inteligenţa, nivelul de fericire sau statutul familial.[24]

Scandalul Cambridge Analytica este un alt episod bine-cunoscut al acestei dezbateri. Firma Cambridge Analytica a fost acuzată de influenţe nelegitime, atât în procesul Brexit, cât şi în alegerile prezidenţiale care au condus la succesul lui Donald Trump. Această interferenţă constă în utilizarea unui set de date obţinute de la Facebook ce permitea prezicerea personalităţii (conform modelului Big Five/OCEAN) a aproximativ 87 de milioane de utilizatori.[25] Pe baza acestor diagnoze ale personalităţii, mai ales a nevrotismului, firma putea personaliza mesajele electorale pentru a le face mai persuasive. De exemplu, un mesaj electoral putea apela la frică pentru a accentua importanţa dreptului de a avea arme, iar un alt mesaj putea apela la tradiţii transmise din tată în fiu, tot pentru a susţine dreptul de a avea arme.[26]

Există şi voci care contestă impactul pe care l-a avut Cambridge Analytica asupra alegerilor prezidenţiale din 2016 sau asupra Brexit. Deşi candidaţii care au apelat la serviciile sale au ieşit câştigători, nu e clar dacă firma a jucat efectiv un rol în rezultatul acestor competiţii. Unele dubii asupra succesului său vin din consideraţii mai generale cu privire la posibila eficienţă a oricărei campanii electorale.

Dată fiind pluralitatea de mesaje şi surse de influenţă de ambele părţi ale dezbaterii, oare cât de mult ar putea influenţa unul, două sau mai multe mesaje? După cum argumenta profesorul de ştiinţe politice Eitan Hirsch în mărturia sa în faţa Comitetului Judiciar al Senatului Statelor Unite, faptul că partidele sau firmele plătesc sume enorme pe campanii publicitare nu înseamnă, automat, că acestea au şi impact.[27]

Totuşi, este greu să invalidăm, de principiu, importanţa marketingului şi posibilitatea sa de influenţare a opiniei publice, dovedită în numeroase instanţe. Hirsch trimite şi la argumente mai specifice, legate de eficienţa metodei alese de Cambridge Analytica pentru ţintirea reclamelor. Mai precis, firma a optat pentru o diagnoză a personalităţii indivizilor pe

modelul Big Five sau OCEAN. Aceasta presupunea inferarea unui tip de personalitate pentru fiecare individ pe baza profilului personal și a prietenilor din Facebook. Mesajele (electorale, publicitare etc.) pe care individul le primea apoi automat erau personalizate în funcție de personalitatea inferată. Variabila-cheie aici este precizia detecției personalității.

Un studiu realizat de David Sumpter pe 20.000 de utilizatori Facebook, al căror profil de personalitate era cunoscut în prealabil, a arătat că precizia detectării corecte este totuși doar puțin mai mare decât simpla presupunere la întâmplare.[28] Așadar, observă autorul, este puțin probabil ca personalizarea reclamelor în funcție de o personalitate estimată aproximativ să fi avut un succes enorm.

Cu toate acestea, analizând controversa legată de Cambridge Analytica și posibilul său impact, Rahul Rathi observă că imprecizia estimării personalității ca predictor electoral nu înseamnă că alte abordări nu pot funcționa foarte eficient.[29] Chiar dacă impactul acestei companii ar fi, de fapt, supraestimat, chiar dacă personalitatea alegătorilor nu ar fi avut cum să fie folosită, în acest caz, pentru personalizarea anunțurilor, Facebook oferă metode de țintire care nu depind deloc de vreo variabilă mediatoare.

Posibilitatea țintirii audiențelor similare se referă la metoda identificării profilului comportamental al celor care votează cu partidul A sau cumpără produsul B, trimițând apoi anunțurile altor indivizi cu un profil asemănător. Datorită vastității datelor disponibile pentru Facebook și brokerilor de date, precum și datorită posibilității rulării unui număr enorm de corelații pentru a identifica predictori eficienți, audiențele pot fi decupate după tot felul de indicatori care semnalizează preferințe și vulnerabilități, chiar dacă nu înțelegem clar cauzalitatea din spate. Este nevoie doar de corelații, nu și de o înțelegere profundă, pentru a putea direcționa anunțuri către un public receptiv. Dacă o combinație de preferințe pentru eroi Marvel *versus* DC, câini *versus* pisici, înghețata de caramel *versus* fructe, tricouri uni *versus* tricouri inscripționate și ciclism versus automobile prezice eficient anumite opțiuni politice, această combinație poate fi utilizată pentru a decupa publicul-țintă al unei reclame electorale fără a înțelege exact rațiunile din spate.

Indiferent de realitatea sau mitologizarea acestor tactici de personalizare și a eficacității lor, scandalul Cambridge Analytica a avut o mare influență în problematizarea etică a utilizării Big Data pentru manipularea la scară largă a alegerilor prin campanii de microțintire electorală.

Există și alte tipuri de manipulări care restrâng autonomia individuală. De exemplu, pe baza datelor personale, o companie poate ținti persoanele dependente de jocuri de noroc pentru a le arăta reclame la cazinouri. Asemănător, o firmă de asigurări poate ținti persoanele care suferă de anumite boli sau regiunile cu o prevalență ridicată a acelor boli pentru a le propune oferte la prețuri dezavantajoase. Totodată, ofertanții de credite cu dobânzi foarte mari pot ținti zonele cu populație vulnerabilă economic și educațional. Astfel de exemple ilustrează puterea asimetrică pe care o au cei care dispun de date *versus* indivizii al căror comportament e influențat și prezis prin date. Big Data și algoritmii folosiți în realizarea de modele și estimări oferă companiilor o putere mare de persuasiune, inclusiv în scopuri exploatatorii.[30]

Nu în ultimul rând, *excesul de precizie are o dimensiune identitară*. Chiar dacă datele par a fi anonimizate, la intersecția unor fluxuri de detalii privind viața indivizilor, devine posibilă asocierea lor cu identitatea unei persoane. În 2006, motorul de căutare AOL a publicat din greșeală pe internet istoricul de căutare pentru 650.000 de utilizatori. Analizele și relatările rezultate în urma acestui incident arată cât de multe informații pot fi inferate și folosite, corect sau incorect, doar pe baza unui set de cuvinte-cheie introduse într-un motor de căutare.

Unul dintre utilizatori a fost identificat nominal de jurnaliștii de la *The New York Times*, care au realizat apoi un articol pe această temă cu consimțământul său.[31] Aflăm, astfel, informații pe un spectru larg de intimitate și chiar de vulnerabilitate despre doamna Thelma Arnold. Aflăm, de asemenea, că unele dintre cuvintele-cheie pe care le-a folosit și unele interese nu se referă la ea, ci la prietenii săi pentru care căuta informații. Se deschide astfel posibilitatea alcătuirii unui profil eronat, dacă ar fi să schițăm un portret automat al tuturor bolilor de care suferă și al planurilor sale de viitor pe baza acestor curiozități.

Investigația realizată de *CNet* pe marginea aceluiași *breach* scoate la

iveală istorii de căutare care pot dezvălui intenţii criminale, experienţe traumatice, comportamente deviante sau chiar ilegale, dar pot fi şi doar semne ale unor curiozităţi psihologice sau literare.[32] Citind această relatare, putem avea experienţa unui deficit de înţelegere, a ambiguităţii inerente oricărei relatări incomplete, dar şi a riscurilor asociate prelucrării automate a unui astfel de set de date.

Scandalul Cambridge Analytica a adus în atenţia publicului un alt aspect al dimensiunii identitare. Trăsăturile, preferinţele şi acţiunile noastre pot fi incluse în propriul profil, fără ca noi să facem absolut nimic în acest sens. Companiile asamblează profiluri din umbră (*shadow profiles*) chiar şi pentru utilizatorii care nu au conturi pe platformele lor, folosind informaţiile date de prietenii şi conexiunile lor. Dacă avem mulţi cunoscuţi cu un cont pe Facebook, este foarte probabil să avem un profil Facebook, chiar dacă nu am avut niciodată un cont pe această platformă.[33]

Ceea ce poate conduce la momente de penibil, la pierderi economice prin furtul de identitate sau la reputaţii ameninţate într-un stat democratic, poate duce şi la pericolul pierderii libertăţii sau vieţii într-un stat autoritar sau totalitar. Exemplul supravegherii continue instalate de guvernul chinez pentru minoritatea uigură[34] este o ilustrare îngrijorătoare a riscurilor excesului de precizie din Big Data. Având la dispoziţie tehnologii de identificare facială tot mai eficiente, inclusiv tehnologii de detectare automate a etniei uigure în funcţie de trăsăturile fizice, statul chinez a implementat politici de control politic şi demografic represiv pentru uiguri şi alte minorităţi, fiind acuzat chiar de genocid.[35] Mărturiile persoanelor care au suferit în urma acestui regim, precum şi o serie de investigaţii jurnalistice, arată cum camerele de supraveghere sunt folosite atât în închisori politice destinate acestor minorităţi, cât şi în localităţi sau case, unde locuirea devine, în fapt, o versiune mai mult sau mai puţin strictă de arest la domiciliu.

Dimensiunea identitară a excesului de precizie este amplificată de progresul tehnologiilor de recunoaştere facială. Astfel de instrumente pot fi folosite atât în scopuri bune, cât şi în scopuri maliţioase, iar uneori graniţa nu este deloc clară. Am văzut că statul chinez este acuzat că apelează la tehnologii sofisticate de recunoaştere facială şi detectare a etniei pentru

politici represive și de epurare etnică. Aparatul polițienesc și de securitate al multor state folosește astfel de soluții, pentru a identifica atât infractori sau teroriști, cât și disidenți sau protestatari. De la an la an, aceste tehnologii devin mai puternice, promițând o identificare tot mai precisă a unui individ prin izolarea sa de milioane sau miliarde de alți indivizi.

În 2021, compania americană Clearview, care oferă tehnologii de recunoaștere facială agențiilor guvernamentale locale și federale, a anunțat amplificarea volumului de imagini individuale recoltate de pe internet (peste zece miliarde) pentru a spori precizia identificării. Clearview afirmă și realizarea unor progrese în identificarea oamenilor care poartă mască sanitară, precum și posibilitatea de identificare în cazul unor poze în care apar fețe încețoșate sau neclare.[36]

Alte companii și laboratoare de cercetare urmăresc intens posibilitatea recunoașterii persoanelor care poartă mască.[37] Guvernele sunt, se asemenea, interesate de beneficiile oferite de această tehnologie. În negocierile actuale privind reglementarea IA în Uniunea Europeană, mai multe țări, precum Franța, Germania, Ungaria, Suedia și Cehia, insistă să nu fie restricționată recunoașterea facială utilizată în scopuri de securitate națională.[38] Viitorul tehnologiilor de recunoaștere facială și de clasificare automată în categorii de risc, inclusiv pe baze rasiale sau etnice, este greu de anticipat și puternic controversat.

Exces de precizie, sau mistificarea unor coincidențe?

Unde se ciocnește, sar și așchii.

Relatările despre precizia ridicată a predicțiilor bazate pe Big Data, învățare automată și inteligență artificială sunt, cel puțin la prima vedere, contradictorii cu relatările privind numeroasele erori aleatorii și sistematice din aceleași sisteme. Unde se află adevărul? Sunt acești algoritmi atât de precișii, încât să ne poată afecta semnificativ autonomia? Sau este acesta un mit menit să susțină legitimitatea proceselor de automatizare, în interesul comercial sau statal (fie real, fie mascat) al celor care folosesc și livrează aceste tehnologii? Problema este dificil de tranșat. Pentru fiecare situație pot fi găsite interpretări alternative.

De exemplu, povestea privind predicția corectă pe care magazinul Target a realizat-o despre graviditatea unei tinere cliente a fost contestată ca fiind o reprezentare incorectă și mitologizantă. După cum observă Fraser într-o discuție dedicată chiar acestui caz, nici una dintre metricile în general folosite pentru a descrie puterea predictivă a unui algoritm (acuratețea, rata de erori fals pozitive, rata de erori fals negative, precizia și capacitatea de redare) nu este prezentată în articol; mai mult, unele ambiguități pun la îndoială chiar relatarea factuală.[39] Este foarte posibil ca această relatare să glorifice o coincidență: foarte multe femei primesc cupoane pentru produse destinate tinerelor mame, iar adolescenta ar fi putut fi una dintre ele. Câte dintre cele care primesc cupoanele nu sunt, de fapt, gravide? Câte dintre clientele Target care sunt gravide nu primesc, de fapt, cupoane? Acestea fiind spuse, vedem cum anumite coincidențe pot fi considerate semne ale unei precizii (extra)ordinare ale predicțiilor algoritmice, susținând utilizarea acestora chiar și atunci când evaluările cantitative ale succesului lor sunt modeste.

În mod asemănător, eficacitatea predictivă a modelelor folosite de Cambridge Analytica a fost pusă la îndoială de autori care au testat puterea lor de diferențiere.[40] Eficiența predictivă a acestor modele este o miză foarte importantă pentru multiple părți implicate în controversă. Scopul cărții de față nu este să adjudece aceste dezbateri, ci doar să observe mizele etice ale discuției.

Dacă predicțiile bazate pe Big Data sunt suficient de precise încât să permită manipularea proceselor electorale și a deciziilor de consum ale indivizilor, ne aflăm într-o situație de degradare a autonomiei individuale, dar și colective, inclusiv de depreciere a instituțiilor democratice. Dacă, dimpotrivă, aceste predicții sunt marcate de erori frecvente și dificil de detectat, procese importante bazate pe decizii automate riscă să încorporeze un grad ridicat de arbitrar și de părtinire, devenind nejustificabile moral.

Adesea, tehnologiile ajung să fie implementate datorită unor promisiuni care aprind imaginația fondatorilor și utilizatorilor. Pentru a avea succes comercial nu e nevoie de o validare științifică a impactului sau de o analiză atentă a posibilităților de eroare sau discriminare. Iluzia obiectivității și eficienței intrinseci a algoritmilor[41] face ca, adesea, să avem mai multă

încredere în decizii automate decât în decizii luate de oameni. Sunt oare reale imparţialitatea şi eficacitatea inteligenţelor artificiale şi ale algoritmilor, sau sunt doar un miraj?

Este important să observăm şi faptul că o controversă se aprinde adesea pornind de la experienţele stranii ale câtorva persoane sau de la declaraţiile cu scop de marketing ale unor companii tech. Raportarea la un studiu sistematic şi ştiinţific sau la o investigaţie jurnalistică de anvergură poate să apară doar ulterior şi cu impact redus. Managerii companiilor de tehnologie au interesul de a supraestima precizia şi eficienţa soluţiilor pe care le vând. Controversele privind Cambridge Analytica[42] sau tehnologiile de recunoaştere facială Clearview[43] ilustrează foarte bine posibilitatea diferenţei dintre promisiuni comerciale şi impact. Experienţele individuale pot fi coincidenţe, într-un volum vast de întâmplări simultane, sau pot fi simptome ale unor tendinţe profunde – dar acest lucru, adesea, nu poate fi decis decât prin studii minuţioase şi de durată.

Exces de accesibilitate şi dreptul de a fi uitat

Apa, vântul şi gura lumii nu o poţi opri.

Dacă vom căuta pe Google numele Mario Costeja González, vom obţine pagini de rezultate în care este detaliat succesul obţinut de cetăţeanul spaniol împotriva gigantului Google, privind aşa-numitul „drept de a fi uitat“. Situaţia era însă complet diferită în 2012, când primele rezultate ale căutării erau nişte anunţuri vechi, din 1998, de scoatere la licitaţie silită a unor proprietăţi ale lui Costeja González, pentru a acoperi datorii neplătite.[44] Supărat din cauza daunelor de imagine pe care aceste informaţii, deja depăşite şi irelevante, i le aduceau, cetăţeanul a solicitat Google să scoată linkurile la aceste anunţuri din rezultatele căutării. Google a refuzat; prin urmare, González a dat compania în judecată, câştigând procesul şi, ulterior, apelul. Datorită cazului său, legislaţia europeană şi, mai târziu, GDPR au clarificat dreptul cetăţenilor de a solicita uitarea informaţiilor inadecvate, irelevante sau excesive – chiar dacă nu sunt neapărat negative.

Practic, nu este efectiv vorba despre un drept de a fi uitat; Carta Europeană a Drepturilor Omului nu include vreo astfel de prevedere. Totuşi,

din dreptul la viaţă privată şi dreptul la protecţia datelor personale rezultă faptul că putem solicita ştergerea referinţelor sau datelor despre noi care nu servesc interesului legitim, comercial sau public de informare. Pentru González, succesul a fost paradoxal. Datorită victoriei sale, istoria sa de insolvenţă este acum larg publicată şi cunoscută pe tot globul. Deşi el a cerut din nou Google eliminarea unor linkuri la articole pe care le considera calomnioase, cererea sa a fost, de data aceasta, respinsă de autoritatea spaniolă pentru protecţia datelor.[45] Prin lupta sa împotriva Google, González a devenit, între timp, o figură publică, fiind considerat erou sau războinic al internetului. Aşadar, relatările trecute despre problemele sale financiare sunt, acum, de interes public, fiind considerate informaţii relevante, şi nu excesive. Prin opoziţia Google la solicitarea sa iniţială, González a fost pus în situaţia de se sacrifica, involuntar, în numele protejării dreptului la viaţă privată.

Mediul digital suferă aşadar de un alt tip de exces specific, şi anume cel de accesibilitate. Informaţiile postate pot fi arhivate şi preluate de alte site-uri, platforme, baze de date etc., devenind perpetuu disponibile celor interesaţi prin intermediul motoarelor de căutare sau al arhivelor de tipul Wayback Machine[10]. Deşi, în viaţa reală, relevanţa acestor informaţii scade odată cu timpul, ele rămân accesibile şi încă mai pot influenţa viaţa celor vizaţi. Problema este cu atât mai acută în cazul informaţiilor care produc daune reputaţiei – de la scandaluri publice la condamnări în justiţie sau declaraţii discriminatorii ce pot fi folosite ca dovadă a unui defect de caracter.

Dacă înainte ecourile peste spaţiu şi timp ale unor astfel de evenimente erau limitate de memoria umană şi de relevanţa publică a persoanei în cauză, astăzi detalii din viaţa a oricărui individ obişnuit pot rămâne online ani la rândul. Datorită acestui exces de accesibilitate, articolul 17 din GDPR[11], denumit şi „dreptul de a fi uitat“, specifică dreptul persoanelor fizice de a solicita ştergerea datelor lor personale. Această ştergere a datelor personale are loc atunci când stocarea acestora nu mai serveşte scopului legitim pentru care au fost colectate sau atunci când subiectul îşi retrage consimţământul pentru colectarea lor, precum şi în alte situaţii prevăzute de lege.

Exces de informații ilegitime

Să stăm strâmb și să judecăm drept.

În anii ’70, orchestrele de top din Statele Unite ale Americii angajau foarte rar femei, fiind compuse aproape integral din bărbați. Treptat, spre finalul anilor ’90, proporția femeilor în orchestre a crescut semnificativ, ajungând la aproape o treime. Cum s-a întâmplat acest lucru? Nu a fost vorba despre politici de acțiune afirmativă. Acestea ar fi presupus sublinierea genului candidaților în luarea deciziei de angajare, prin acordarea unei priorități mai mari muzicienelor față de muzicieni.

Dimpotrivă, soluția a constat în ascunderea genului, prin așa-numitele audiții în orb (*blind auditions*).[46] Comitetele de selecție ezitau să angajeze femei datorită unor stereotipuri negative, conștiente sau nu. Atunci când candidații și candidatele își demonstrau talentul muzical în spatele unei cortine, stereotipurile de gen puteau fi evitate. Așa au ajuns să fie selectate mai multe femei, pornind doar de la ce se auzea, nu și de la ce se vedea. Prin urmare, o informație ilegitimă, care distorsiona evaluarea meritelor muzicale, a putut fi ascunsă de cei care luau decizia.

În ziua de azi însă mai este oare posibil să ascundem vreo informație de algoritmii ce ne decid soarta CV-urilor sau candidaturilor? Chiar dacă genul sau vârsta nu sunt comunicate în dosar, ele sunt accesibile prin profilurile noastre digitale sau pot fi inferate probabilistic din multe alte informații pe care le-am oferit. Chiar și audițiile în orb au suferit, la început, de problema dezvăluirii involuntare a genului. Muzicienele care intrau pe scenă pășeau cu un zgomot distinctiv, din cauza pantofilor cu toc. Orchestrele au început să solicite candidaților să intre desculți. În lumea digitală este însă imposibil să intrăm desculți, deoarece datele noastre au fost deja arhivate și clasificate în profiluri digitale complexe.

În ianuarie 2021, guvernul olandez a demisionat în urma unui scandal de anvergură privind persecuția a peste 25.000 de familii acuzate, pe nedrept, de fraudă fiscală. Un tată de 46 de ani, Roger Derikx, a fost silit să returneze 68.000 de euro primiți, în mod legitim, ca ajutor pentru copii. Datorită dificultăților financiare enorme cu care s-a confruntat pentru a plăti aceste sume, căsnicia sa s-a destrămat, ajungând la divorț.[47] Alți părinți și-

au pierdut dreptul de a-şi creşte copiii datorită sărăciei în care au fost împinşi.[48] Protestele victimelor, începute încă din 2012, au fost ignorate timp de ani de zile. Cum s-a ajuns aici? Un sistem computerizat calculase riscul de fraudă pentru familiile defavorizate care primeau ajutoare pentru copii, ţinând cont, printre altele, de cetăţenia acestora. Această informaţie este discriminatorie şi ilegitimă în detectarea fraudei, după cum a conchis şi Autoritatea olandeză pentru protecţia datelor.[49] Desigur, mai mulţi factori administrativi şi politici au contribuit la un eşec de asemenea amploare. Remarcăm însă şi cum încrederea prea mare a guvernului olandez în obiectivitatea unui sistem informatic a condus la negarea dreptului victimelor de a-şi exprima protestul şi de a-şi prezenta situaţia în tribunal, prelungind şi amplificând suferinţa acestora şi a familiilor lor.

Aceste exemple ilustrează faptul că un tip problematic de exces al cunoaşterii se referă la disponibilitatea unor date corecte factual, dar incorecte din punct de vedere etic pentru o anumită decizie. Mai precis, produsele de predicţie pe baza Big Data fac inevitabilă disponibilitatea informaţiilor privind genul, vârsta, identitatea etnică sau rasială, confesiunea religioasă, orientarea sexuală, cetăţenia etc. – informaţii ce pot deveni o sursă de discriminare.

În lumea offline, este considerat ilegitim şi chiar ilegal ca procesele decizionale să ţină cont de astfel de trăsături ale candidaţilor. De exemplu, articolul 21 din Carta Europeană a Drepturilor Omului specifică: „Se interzice discriminarea de orice fel, bazată pe motive precum sexul, rasa, culoarea, originea etnică sau socială, caracteristicile genetice, limba, religia sau convingerile, opiniile politice sau de orice altă natură, apartenenţa la o minoritate naţională, averea, naşterea, un handicap, vârsta sau orientarea sexuală“. Prin urmare, organizaţiile trebuie să ignore aceste trăsături atunci când iau decizii pentru care astfel de informaţii nu sunt relevante. Ignoranţa, sau cel puţin ignorarea, devine necesară pentru decizii virtuoase.

Un exemplu de asumare a ignoranţei pentru a înlesni o decizie etică se regăseşte în concursurile organizate pentru angajarea muzicienilor în orchestre. Aşa cum am menţionat mai sus, date fiind riscurile de discriminare pe bază de gen şi rasă, au fost instaurate practici de „audiţii în orb“, în care candidaţii îşi arată talentul în spatele unui paravan. Comisia de

evaluare are astfel acces doar la aspectele considerate relevante pentru munca lor, fiind pusă în situația de a ignora ceea ce, actualmente, este considerat irelevant.[50] În contextul deciziilor de recrutare care folosesc soluții de IA pentru evaluarea oamenilor, o astfel de ignorare devine imposibilă. Trăsăturile socio-demografice sunt adesea *incluse explicit* în fluxurile de date publice arhivate și combinate în profilurile noastre digitale. Chiar dacă nu sunt, ele vor fi *inferate* din alte informații corelate. O soluție de inteligență artificială care optimizează în funcție de un set de criterii ce pare neutru în funcție de gen, vârstă etc. poate învăța că aceste trăsături sunt predictori utili și îi poate *factoriza implicit*, chiar fără ca dezvoltatorii umani sa fie conștienți de acest lucru.[51] Cu atât mai mult, dacă o organizație este interesată de aceste trăsături, pentru construirea unui profil în scop politic, militar sau comercial, aceste date sunt ușor accesibile sau pasibile de a fi inferate din profilurile de date personale. Prin urmare, devin necesare procese explicite de testare e echității și de monitorizare a discriminării pe baza unui set de criterii considerate potențiale riscuri.[52]

Sistemele de intervenție predictivă (*predictive policing*) din domeniul ordinii publice oferă un exemplu de riscuri ce derivă din excesul de informații ilegitime.[53] Aceste sisteme realizează profiluri ale zonelor cu risc ridicat de criminalitate, precum și profiluri pentru școli sau indivizi, estimând riscuri și probabilități ale unor acte contrare legii. Estimările acestor probabilități, materializate în grafice sau hărți de tip *heatmaps*, devin apoi reper pentru intervenții preventive ale forțelor de ordine.

Există însă cel puțin două tipuri de efecte nocive, de restrângere a autonomiei individuale, care pot deriva din aceste acțiuni.

Prin *efectul de etichetare* (*labeling theory*), dacă o persoană este sistematic etichetată ca fiind probabil deviantă și cu risc ridicat de violență, această predicție s-ar putea împlini de la sine. Indivizii reacționează la definițiile pe care ceilalți le au despre ei, iar aceste definiții modifică structurile de oportunități, creând uneori chiar fenomenele negative pe care încearcă să le prevină.

Efectul de părtinire confirmatoare (*confirmation bias*) apare atunci când luăm în considerare cu o mai mare intensitate zonele în care suspectăm o incidență mai ridicată a unui fenomen, ceea ce perpetuează percepția

conform căreia un fenomen se petrece mai mult într-un loc decât în altul. De exemplu, dacă forţele de ordine sunt concentrate în zonele definite ca fiind rău-famate şi cu risc major de criminalitate, unde efectuează mai multe controale şi intervenţii, este posibil ca lipsa lor din alte zone, unde criminalitatea rămâne nedetectată, să conducă, parţial, la reproducerea acestor două tipuri de efecte.

Pentru a preîntâmpina astfel de probleme, Asaro[54] propune două abordări etice diferite cu scopul de a înţelege şi orienta intervenţiile poliţieneşti pe baza Big Data. Pe de o parte, el identifică aşa-numitul model al ameninţărilor (*threat model*), tipic intervenţiilor şi politicilor care definesc indivizii ca fiind posibile ameninţări cu grad mai mare sau mai mic de risc. Acest model încurajează intervenţii represive, cu o definiţie negativă a situaţiei privind oamenii vizaţi. Autorul contrastează modelul ameninţărilor cu cel al eticii grijii (*ethics of care*), în care Big Data sunt folosite pe două paliere. Primul nivel vizează utilizarea Big Data în diagnoza riscurilor de devianţă şi al tensiunilor sociale, iar al doilea are în vedere intervenţiile de soluţionare bazate pe remedierea surselor acestor probleme, precum şomajul sau sărăcia.

Prin urmare, riscurile de discriminare derivând din excesul de informaţii ilegitime apar atunci când diagnozele statistice sunt cuplate cu acţiuni represive sau de restrângere a drepturilor individuale, aceste riscuri putând fi reduse prin acţiuni de remediere şi creare de oportunităţi.

Excesul costurilor cunoaşterii

Ai, dai, n-ai. Ia nu da, să vezi cum ai.

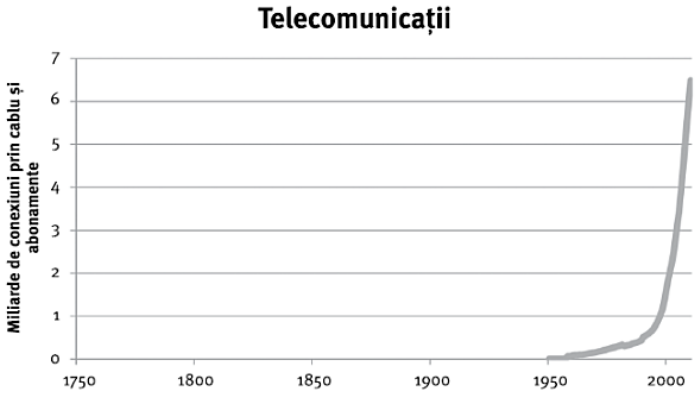
Un alt tip de exces derivat din Big Data se referă la costurile cunoaşterii, mai precis, la costurile financiare de antrenament pentru maşinile de învăţare automată şi inteligenţă artificială care procesează Big Data.

În ultimii ani, tipurile de soluţii ML practicate la frontiera tehnologiei, de către marii giganţi tehnologici, precum Google, Amazon sau Facebook, au crescut în complexitate, presupunând volume tot mai ample de date şi o putere de calcul tot mai mare. Estimările de bugete pentru antrenamentul unor astfel de modele avansate sunt de zeci de mii de dolari[55], sau chiar de

milioane de dolari[56]. Aceste costuri financiare accesibile doar marilor corporații pot conduce la consolidarea unei asimetrii de putere deja existente între platformele digitale și utilizatorii lor. În plus, poate avea loc și distorsionarea mai accentuată a pieței și a competiției, restrângând autonomia indivizilor și a firmelor mici în promovarea propriilor idei și interese.

Un reportaj *MIT Technology Review* observă că nivelul tot mai ridicat al costurilor conduce și la privatizarea cercetării în domeniul tehnologiei, în contextul în care domeniul academic nu are la îndemână astfel de bugete: „Simpla dimensiune a resurselor necesare azi pentru a produce rezultate demne de publicare ridică provocări serioase în calea contribuțiilor cercetătorilor din domeniul academic“[57]. Pot apărea astfel riscuri de scădere a transparenței și de concentrare a cunoașterii, cu potențiale efecte negative asupra monitorizării costurilor și riscurilor acesteia.

Costurile financiare reflectă și impactul ecologic al tehnologiilor. Economia globală este definită de bucle de feedback pozitiv prin care tehnologiile informației și producția de bunuri și servicii se amplifică reciproc, ducând, indirect, la creșterea consumului de energie și a emisiilor de CO_2 (vezi figura 32). Amprenta de carbon a modelelor tot mai complexe a crescut neliniar, ajungând la praguri considerate îngrijorătoare, mai ales având în vedere tendințele de accelerare și consolidare a domeniului tehnologiei informației.[58] În același timp, *eficiența energetică a centrelor de date a crescut semnificativ în ultimii ani*, ceea ce reprezintă un motiv de optimism privind alocarea etică a resurselor și potențialul impact asupra mediului.



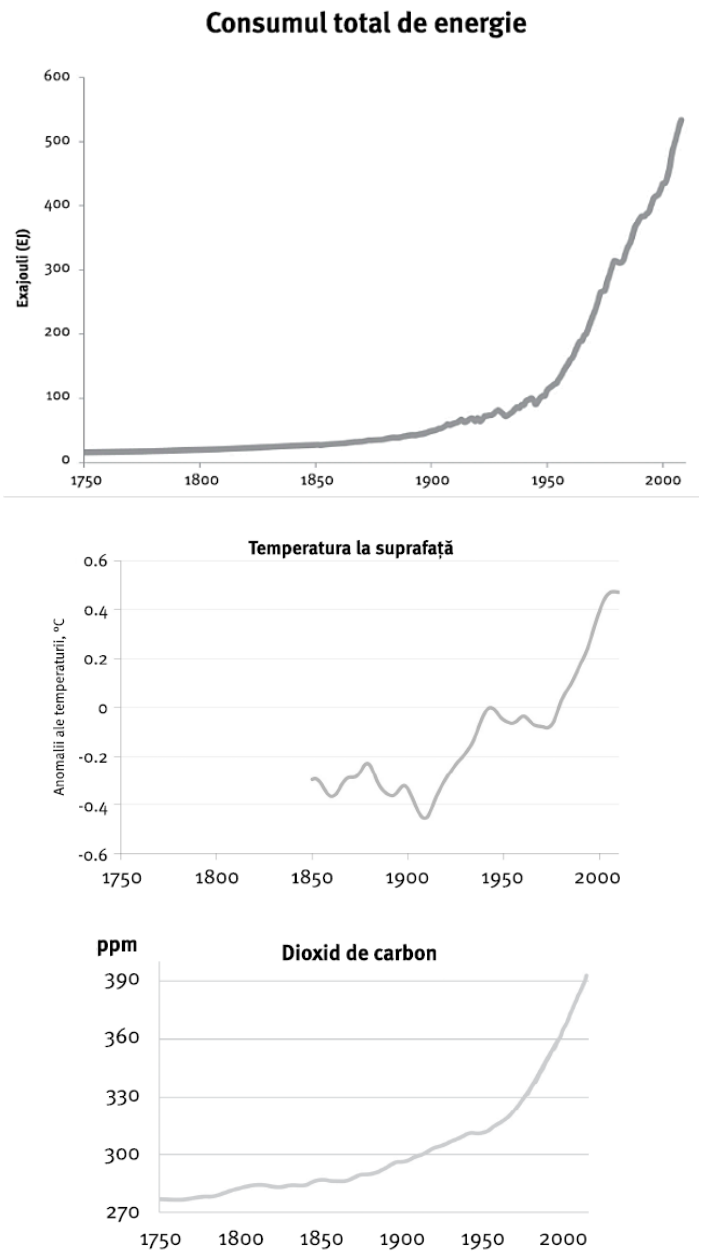


Fig. 32 Tendințe de creștere accelerată și interconectată privind telecomunicațiile, consumul total de energie, emisiile de dioxid de carbon și creșterea temperaturii terestre. Sursa datelor și a graficelor: IGBP – International Geosphere Biosphere Programme.

Un studiu publicat în revista *Science* a ajuns la concluzia că estimările anterioare privind impactul energetic global al Big Data au fost exagerate, înregistrându-se rezultate mult mai bune decât cele prezise inițial. Dacă, în ansamblu, consumul total de energie al centrelor de date a crescut cu 6% în intervalul 2010–2018, atingând aproximativ 1% din consumul global de

energie, totuşi, în aceeaşi perioadă, volumul sarcinilor realizate în centrele de date a crescut cu 550% – ceea ce, în acest interval de timp, reprezintă o scădere a consumului de energie per sarcină realizată cu 20% anual.[59] Cu toate astea, este neclar în ce măsură această decuplare relativă (rata de creştere a resurselor utilizate este inferioară ratei de creştere economică) va conduce la o decuplare absolută (creştere economică sporită şi bunăstare cu o utilizare mai redusă a resurselor naturale şi impact diminuat asupra mediului)[60] în privinţa intensificării activităţilor digitale în raport cu creşterea emisiilor de carbon şi a consumului de resurse terestre.[61]

Nu în ultimul rând, *vastele volume de date digitale au şi un cost psihologic*, suportat de un segment ocupaţional distinct: aşa-numiţii *moderatori de conţinut*, care sunt responsabili cu filtrarea postărilor obscene, abuzive şi/sau violente. Această categorie profesională este expusă unor condiţii foarte dificile, chiar traumatizante, de muncă. Astfel de angajaţi sunt izolaţi unii de alţii, precum şi de cei apropiaţi lor, prin clauzele contractuale care le impun să nu dezvăluie detaliile propriei activităţi.

Reportaje multiple, precum cel publicat în *Wired*[62] sau în *The Verge*[63], privind condiţiile de muncă şi traumele la care sunt expuşi aceşti oameni, conduc la problematizarea costului asimetric pe care societatea îl plăteşte pentru a beneficia de internetul colaborativ şi pentru a posta vastele cantităţi de mesaje ce sunt parte din Big Data, dar care trebuie să respecte anumite rigori calitative şi, implicit, etice.

Excesul relativ: (dez)echilibrul cunoaşterii

Sătulul nu crede flămândului.

În 2016, aplicaţia de dating Grindr, folosită mai ales pentru relaţii homosexuale, a fost achiziţionată de firma chineză Beijing Kunlun Tech Co Ltd. Trei ani mai târziu, când compania se pregătea de listare la bursă, guvernul Statelor Unite a decis că această aplicaţie nu se mai poate afla în proprietatea unei firme chineze, din cauza riscurilor de securitate naţională.[64]

În 2020, Kunlun fost silită să vândă Grindr unei companii americane, San Vicente Acquisition Partners, avizată de autorităţile relevante (şi anume

Comitetul pentru Investiţii Străine din Statele Unite).[65] Cum se explică această intervenţie dură a statului american împotriva unui investitor, în lipsa oricărui scandal public?

Grindr aduna şi stoca informaţii extrem de sensibile despre cei 27 de milioane de utilizatori ai săi: de la conversaţii intime şi poze erotice până la poziţionarea geografică detectată la nivel de metri. Dat fiind că şi personalul militar al SUA foloseşte aplicaţia, precum şi alţi indivizi cu roluri importante în politică sau serviciile de informaţii, toate aceste date îi făceau ţinte ale şantajului. Statul american a înţeles vulnerabilitatea la care se expune prin accesul unei companii chineze, potenţial controlate de statul chinez, la astfel de date.

În acelaşi timp, utilizarea informaţiilor din Grindr de către terţe părţi americane interesate împotriva unor cetăţeni americani care nu sunt legaţi de armată sau de serviciile de informaţii a rămas în continuare posibilă şi complet legală. O dovadă este cazul unui important preot catolic denunţat de *The Pillar*, o publicaţie care a cumpărat date de la Grindr, ca fiind homosexual[66], ceea ce a condus la demisia sa şi la un scandal de proporţii. Investigaţia *The Pillar* a scos la iveală şi alte cazuri în care aplicaţia Grindr a fost folosită în reşedinţe ale clericilor catolici (fără a încerca, de data asta, să arate originea datelor şi să identifice utilizatorii) atât din Vatican,[67] cât şi din SUA[68]. De altfel, autoritatea norvegiană de protecţie a consumatorului a amendat Grindr cu 7,16 milioane de euro pentru vânzarea datelor utilizatorilor fără consimţământul acestora.[69] La începutul acestui capitol, am văzut că numeroase dispozitive smart, precum glucometrele sau monitoarele cardiace, transmit datele despre sănătatea utilizatorilor către companiile de asigurări, adesea fără consimţământul real şi informat al clienţilor – putând duce la reducerea accesului acestora la serviciile medicale. Nenumărate aplicaţii din telefonul mobil obţin date privind localizarea noastră, actualizată de mii de ori pe zi cu o înaltă precizie, permiţând astfel înţelegerea clară a locurilor în care ne petrecem timpul. Aceste date sunt apoi vândute pe o piaţă foarte lucrativă, în care informaţiile privind deplasările noastre devin indicii utile despre valoarea pe care o avem ca ţinte ale diferitelor negocieri, reclame sau oferte comerciale.[70] În 2020, o investigaţie *Motherboard* a arătat că armata

Statelor Unite cumpăra datele de localizare geografică vândute de o aplicație de rugăciuni destinată credincioșilor musulmani, precum și de la o aplicație de dating orientată spre musulmani.[71] O investigație a *The Wall Street Journal* a arătat cum guvernul SUA folosește datele de localizare obținute prin aplicații meteo, de comerț online sau jocuri de mobil, pentru controlul populației și al imigranților[72], în ciuda unei decizii anterioare a Curții Supreme de Justiție, care prevede necesitatea unui mandat judecătoresc pentru a accesa datele de localizare din telefonul unei persoane.[73] Este tot mai greu să știm cu adevărat cine anume are acces la istoricul detaliat al localizărilor înregistrate de telefonul nostru mobil și ce face cu acest istoric. Circuitul datelor dinspre aplicații și dispozitive spre companii care le vând apoi unor terțe părți interesate deschide noi și noi posibilități de atac împotriva indivizilor – în numele moralității, profitului, răzbunării personale, securității naționale sau a oricărui interes susținut de banii necesari pentru a le cumpăra.

Problematica excesului de cunoaștere este amplificată de problematica dezechilibrului acesteia. Dezechilibrul are două dimensiuni: decontextualizare și asimetrie informațională. În primul rând, *fluxurile masive de date alterează logica contextuală a informațiilor*[74], scoțându-le din cadrul interacțional în care acestea dădeau avantaje reciproce participanților și agregându-le într-un spațiu în care sunt disponibile oricui poate să le cumpere. Dacă am dat acces la datele mele de localizare pentru a afla cum e vremea mâine, când trebuie să merg la un control medical, nu înseamnă că sunt de acord ca aceste date să fie folosite pentru a-mi trimite reclame la medicamente. Dacă am dat acces la date privind înfățișarea mea, deoarece vreau să experimentez cu noi forme de tunsoare și barbă fără să merg la frizer, nu înseamnă că sunt de acord ca fața mea să fie folosită pentru antrenarea algoritmilor de recunoaștere facială utilizați apoi, în scop militar, de SUA, China sau oricine le-a cumpărat. Decontextualizarea este și una dintre sursele deficitului de înțelepciune practică, sau *phronesis*, în deciziile bazate pe Big Data, așa cum vom discuta și în capitolul următor. Datele sunt agregate în ansambluri care nu mai fac posibilă menținerea nuanțelor, ponderilor, structurilor de relevanță sau irelevanță ale relațiilor și contextelor în care au fost generate.

În al doilea rând, *accesul la date și controlul acestora este profund inegal*. Achiziționarea și prelucrarea Big Data presupune resurse financiare, computaționale și tehnologice semnificative, precum și expertiză și forță de muncă. Marii giganți tehnologici și alte mari organizații sunt privilegiate în accesul la aceste resurse, putând să le folosească pentru a evalua, clasifica sau anticipa comportamentul celor ce au generat datele. După cum observă Andrejevic[75], distincția între bogații și săracii datelor devine o distincție între cei care sortează și cei care sunt sortați, între subiecții deciziilor și obiectele deciziilor, consolidând asimetrii de putere.

Dacă, anterior Big Data, accentul se punea pe fractura digitală (*digital divide*), creșterea considerabilă a accesului la internet în ultimele decenii a condus la o nuanțare a conceptului și a indicatorilor. Astăzi, decalajul digital se conjugă cu decalajul accesului la date, care însă nu are loc doar între diferite clase sociale de indivizi sau între cetățeni ai unor țări cu niveluri de dezvoltare diferite. Decalajul accesului la date diferențiază și indivizii și organizațiile, respectiv persoanele fizice – ale căror date sunt preluate, stocate și modelate – de persoanele juridice – care preiau, stochează, modelează și utilizează datele în scop predictiv. Desigur, oamenii implicați în aceste organizații vor beneficia de pe urma accesului la date, în scop comercial și politic, în timp ce oamenii cu legături mai slabe la fluxurile de date și beneficiile lor vor fi mai degrabă resursele decât câștigătorii acestui proces.

Nu doar că indivizii nu au puterea de a colecta și procesa Big Data, dar, adesea, nu au nici acces la cunoașterea proceselor în care Big Data sunt folosite și a influenței acestora asupra vieții lor. Termenii și condițiile de utilizare oferite de furnizorii de produse și servicii digitale utilizatorilor lor sunt într-atât de complecși, încât o persoană tipică le va accepta, de cele mai multe ori, fără o inspectare atentă.[76]

Nu este vorba doar despre încercarea de a influența utilizatorii să accepte o deschidere cât mai mare spre a ceda date – deși și această manipulare comercială are un rol important. Mai mult, chiar dacă o companie ar dori să fie transparentă și să specifice clar și detaliat procesele de colectare și utilizare a datelor, acest lucru ar fi foarte dificil de realizat efectiv, deoarece, într-o mare măsură, aceasta transferă datele (și) unor terțe părți, precum

conglomeratele din care face parte, sau către *data brokers*, cu care are relaţii comerciale. Circuitele şi utilizările fluxurilor de date sunt adesea diverse, fluide şi complexe, nefiind inteligibile clienţilor obişnuiţi.

Prin urmare, exigenţa transparenţei utilizării datelor este dificil de atins în practică. De aici reiese importanţa implicării autorităţilor publice, precum cele stabilite de GDPR, şi a altor organizaţii cu rol de monitorizare şi sprijin în favoarea indivizilor.

Chiar şi atunci când legea înclină balanţa înspre cetăţeni, precum GDPR, este în continuare dificil să obţinem accesul la datele noastre şi să decidem ce să facem cu ele odată ce le-am obţinut. Cei care reuşesc să solicite marilor companii care ne însoţesc viaţa, precum Amazon, Facebook sau Tinder, să le elibereze o copie a tuturor datelor lor personale, sunt adesea uimiţi de nivelul de detaliu al înregistrărilor. Când Alina Utrata, studentă doctorandă la Universitatea Cambridge, a obţinut de la Amazon datele sale personale, a constatat că acestea includeau fişierele audio ale tuturor conversaţiilor sale cu Alexa, fişiere Excel cu toate cumpărăturile sale de pe Amazon sau WholeFoods, detalii privind ce şi cum citeşte pe Kindle şi toate vizitele sale la magazinele fizice Amazon.[77] Frederike Kaltheuner şi-a obţinut profilul digital de la brokerul de date Quantcast, combinând date detaliate despre cumpărăturile sale, inferenţe despre situaţia familială şi financiară, preferinţe, călătorii şi tot felul de clasificări şi etichetări pe baza plăţilor făcute cu cardul şi a site-urilor vizitate.[78] Deşi, teoretic, ştim că tot ce facem este stocat online şi transmis mai departe, este uneori dificil să ne imaginăm ce înseamnă acest „tot“. Dacă totuşi vrem să trecem la acţiune, cum putem să limităm accesul platformelor şi brokerilor de date care ne înregistrează de ani de zile? Pare o muncă de Sisif, o luptă cu morile de vânt, mai ales dacă vrem totuşi să continuăm să avem o prezenţă digitală. Există speranţa că acest efort poate fi, cel puţin parţial, delegat unor boţi, pe principiul cui pe cui se scoate. Cel puţin, aceasta este promisiunea unui startup din California, unul dintre puţinele state din SUA care au o lege privind protecţia datelor personale. Totuşi, consumatorii care doresc să-şi şteargă profilul acumulat de brokerii de date ar trebui să solicite ştergerea datelor lor de la fiecare firmă, rând pe rând, ceea ce reprezintă un efort considerabil. Startup-ul DoNotPay oferă serviciul de a solicita automat

ștergerea profilurilor digitale de la peste 100 de brokeri de date[79], contra unei taxe lunare de 3 dolari. Pentru cetățenii statelor fără o astfel de legislație, acest ajutor rămâne inutil. Riscăm să fim, așadar, rătăciți în propriile urme digitale, pe care alții par că le decodifică mai bine decât noi, și nu neapărat în favoarea noastră. Adesea, nu ne este foarte clar ce instrumente sau ce parteneri ne pot veni în ajutor și cum putem înclina balanța în favoarea noastră.

Prin urmare, este nevoie de o bună (re)definire a regulilor jocului la nivel social, combinată cu sprijin pentru indivizi, pentru a conduce la restabilirea echilibrului de putere între cetățeni și organizațiile care preiau, stochează, procesează și folosesc datele noastre personale.

---

10. Disponibil online la https://archive.org/web.
11. Disponibil online la https://gdpr-info.eu/art-17-gdpr.

4. Deficitul cunoașterii în societatea algoritmică

Prostul dacă nu-i fudul, parcă nu e prost destul.

Am văzut în capitolul anterior că abundența datelor se poate transforma în exces, mai ales atunci când devine resursă într-un joc de putere asimetric. Informații extrem de detaliate despre acțiunile noastre sunt înregistrate și transmise, pentru profit, către brokeri de date care le dau mai departe, cel mai adesea fără să fim conștienți de amploarea, destinația sau utilizarea finală a acestor fluxuri informaționale. Organizații comerciale, politice sau publice pot lua decizii în ceea ce ne privește pe baza unor indicii despre noi bazate pe date asupra cărora nu mai avem nici un control. Dezanonimizarea prin excesul de precizie, discriminarea algoritmică și dezavantajele în negocierile cu cei care ne vor banii sunt câteva dintre riscurile la care suntem expuși în noul capitalism al datelor. La nivel social, tot mai multă putere este acumulată de către cei ce dețin deja puterea computațională și informațională. Marii giganți comerciali și statali își consolidează poziția prin combinarea datelor vaste cu modele de inteligență artificială antrenate și perfecționate pe aceste enorme colecții de informație. Aceștia generează modele masive de control al populațiilor, care nu mai pot fi cu adevărat înțelese sau testate sistematic din exterior, ci doar observate. Totuși, atunci când examinăm riscurile guvernanței algoritmice, aceasta este doar o fațetă a situației. Reversul medaliei se referă tocmai la limitele, erorile și lipsurile cunoașterii generate de Big Data și învățarea automată. Circuitele datelor par să genereze un telefon fără fir, în care multe mesaje ajung grav distorsionate. Câți dintre noi riscăm să cădem printre crăpăturile noului edificiu informațional?

Prin urmare, o a doua direcție de analiză a eticii deciziilor algoritmice și Big Data, prin raportare la cunoașterea pe care o generează, constă în studiul deficiențelor produselor bazate pe date vaste. În ce măsură modelele algoritmice bazate pe Big Data produc reprezentări corecte și adevărate ale lumii? Care sunt limitele acestor reprezentări? Care sunt consecințele

acţiunilor pe care organizaţiile le întreprind pe baza acestor reprezentări şi a erorilor aferente? Care este relevanţa morală a acestor acţiuni?

De exemplu, în 2021, Facebook avea aproape trei miliarde de utilizatori, dintre care două treimi foloseau o altă limbă decât engleza. Mai precis, utilizatorii Facebook folosesc peste 110 limbi pe platformă. Totuşi, moderarea conţinutului se realizează cel mai intens şi mai competent pentru limba engleză. Chiar şi aici, ştirile false şi polarizarea reprezintă probleme semnificative.[1] Pentru alte limbi, precum araba şi dialectele sale, erorile de detectare a mesajelor problematice par să fie considerabile.

Un document intern al companiei, dezvăluit în aşa-numitele Facebook Files[2] de către avertizoarea de integritate Frances Haugen, menţiona că avertismentele de conţinut terorist pentru limba arabă sunt greşite în 77% din cazuri[3] (deşi compania a disputat această estimare). Problema apare şi mai mult pentru limbile pentru care inteligenţa artificială nu poate oferi traduceri nuanţate, precum cele vorbite în multe state sărace ale globului. În Sri Lanka, guvernul a blocat Facebook în repetate rânduri, considerând că nu face faţă barajului de incitări la ură.[4] În Myanmar, Facebook este acuzat că a permis intensificarea violenţei genocidale împotriva minorităţii musulmane rohingya, datorită eşecului de a modera conţinutul în limbile locale.[5]

După cum sugerează şi proverbul românesc citat în mottoul acestui capitol, deficitul de cunoaştere este şi mai condamnabil atunci când trece drept înţelepciune. Proverbul ilustrează destul de bine criticile aduse soluţiilor bazate pe Big Data şi inteligenţă artificială. Acestea sunt simultan fudule, adică lăudate ca o întruchipare a unor forme avansate de cunoaştere, şi manifestări ale prostiei, fiind criticată lipsa de inteligibilitate umană a modelelor bazate pe învăţare automată.

Putem ilustra această tensiune dintre înaltele promisiuni privind cunoaşterea livrată şi riscurile sau limitările Big Data – de asemenea, semnificative ca impact – prin postarea lui Tom Shafer, „The 42 V’s of Big Data and Data Science“ („Cei 42 de V ai datelor masive şi ai ştiinţei datelor“)[6]. Schafer reacţionează, parţial umoristic, parţial cu spirit investigativ, la creşterea continuă a numărului de litere V folosite pentru abrevierea specificităţilor Big Data.

Încă din 2001, datele vaste au fost caracterizate prin „cei trei V“, o sintagmă devenită celebră. Aceşti primi trei V erau: volum, varietate şi viteză. Volumul se referă la cantitatea imensă de date specifică acestui tip de modelare, trăsătură distinctivă invocată şi în conceptul de Big Data. Varietatea se referă la eterogenitatea informaţiilor care au formatări diferite şi nu sunt structurate după un tipar unitar, aşa cum erau datele folosite anterior de statisticieni. Big Data pot combina inclusiv date numerice, text, reţele, audio-video etc. Viteza surprinde ritmul accelerat în care datele sunt generate şi acumulate în societatea actuală.

Zece ani mai târziu, în urma consolidării utilizării Big Data, apare un nou V distinctiv, şi anume valoarea pe care Big Data promit să o aducă organizaţiilor. Numărul de V-uri începe să crească, observă Schafer, adăugându-se apoi veracitatea. Veracitatea este prima menţiune explicită a unei rezerve privind valoarea de cunoaştere: sunt aceste date veridice? Aceşti 5V ai Big Data (volum, varietate, viteză, valoare şi veracitate) reprezintă una dintre descrierile frecvent folosite actualmente. Ulterior, după cum putem vedea în prezentarea lui Schafer, şi-au mai făcut apariţia alte două trăsături: volatilitatea (tendinţa Big Data de a se schimba impredictibil, inclusiv prin manifestarea unor erori) şi validitatea (problematizarea relevanţei modelelor bazate pe Big Data pentru rezolvarea unor probleme actuale).

Postarea lui Schafer este interesantă datorită celor 42 de V pe care autorul îi propune şi pe care i-am clasificat în tabelul de mai jos în funcţie de tonul lor optimist sau pesimist privind valoarea cunoaşterii. Dintre cele 42 de trăsături menţionate de Schafer, 20 sunt optimiste, 5 sunt pesimiste, iar 17 sunt neutre, nefiind relevante pentru valoarea de adevăr. Observăm şi în acest caz tonul general pozitiv al dezbaterii publice, precum şi promisiunile ample făcute în numele Big Data, şi anume de a aborda problemele mari ale omenirii (*Valor* şi *Vexed*), de a trece în spatele vălului aparenţelor prin identificarea cauzelor ascunse (*Veil*), de a oferi cunoaştere valoroasă uşor accesibilă (*Vanilla*), de a testa asumpţii (*Vet*), de a combina perspectiva de ansamblu cu cea de detaliu (*Varifocal*), de a oferi cunoaştere acţionabilă (*Vivify*, *Vocabulary* şi *Voice*) şi de a crea o comunitate de cunoaştere (*Vibrant*).

Cele cinci limitări menţionate de autor se referă la riscul erorilor din date (*Veracity* şi *Volatility*) şi din software (*Varmint*), la ambiguitatea ireductibilă a concluziilor trase pe baza tiparelor din Big Data (*Vagueness*) şi la dificultatea traducerii datelor în modele relevante pentru lumea reală (*Viability*).

Tabelul următor arată conotaţiile pozitive şi negative ale specificităţilor atribuite Big Data, stabilite în urma unei analize proprii a trăsăturilor Big Data menţionate în postarea lui Tom Shafer.[7]

V-uri specifice	Trăsătura specifică Big Data (ro.)	Atitudine privind cunoaşterea
1. *Validity*	Validitate: Rigoarea în analiză este esențială pentru predicții valide.	Optimistă
2. *Valor*	Îndrăzneală: În ciuda (sau tocmai datorită) datelor vaste, trebuie să abordăm cu brio marile probleme.	Optimistă
3. *Value*	Valoare: Știința datelor continuă să ofere o valoare din ce în ce mai mare pentru utilizatori, pe măsură ce tot mai multe date devin disponibile și sunt dezvoltate noi tehnici.	Optimistă
4. *Vane*	Giruetă: Știința datelor poate indica direcția corectitudinii procesului de luare a deciziilor.	Optimistă
5. *Vanilla*	Simplitate: Chiar și cele mai simple modele, construite cu rigoare, pot oferi valoare.	Optimistă
6. *Vantage*	Avantaj: Datele vaste ne permit să avem un punct de vedere privilegiat asupra sistemelor complexe.	Optimistă
7. *Varifocal*	Varifocal: Big Data și știința datelor ne permit să vedem atât pădurea, cât și copacii.	Optimistă
8. *Vaticination*	Viziune: Analiza predictivă oferă capacitatea de a prognoza. (Desigur, aceste previziuni pot fi mai mult sau mai puțin precise, în funcție de rigoarea și de complexitatea problemei.)	Optimistă
9. *Veer*	Viraj: Odată cu dezvoltarea științei datelor agile, ar trebui să fim capabili să navigăm printre nevoile clientului și să schimbăm direcțiile rapid atunci când situaţiile o cer.	Optimistă
10. *Veil*	Văl: Știința datelor ne oferă capacitatea de a privi în spatele cortinei și de a examina efectele variabilelor latente din date.	Optimistă
11. *Versed*	Versat: Experţii în știința datelor trebuie adesea să știe câte puțin din fiecare: matematică, statistică, programare, baze de date etc.	Optimistă
12. *Vet*	Verificare: Știința datelor ne permite să ne examinăm presupunerile, contribuind cu dovezi la valoarea intuiției.	Optimistă

13. *Vexed*	Dispută: O parte din entuziasmul din jurul științei datelor se bazează pe potențialul său de a face lumină asupra unor probleme mari și complicate.	Optimistă
14. *Vibrant*	Vibrant: O comunitate înfloritoare de experți în știința datelor este vitală și ne oferă perspective, idei și sprijin în toate eforturile noastre.	Optimistă
15. *Virtuosity*	Virtuozitate: Deși experții în știința datelor trebuie să știe doar câte puțin din foarte multe domenii, este necesar să continuăm să aprofundăm fiecare domeniu, înțelegând tot mai multe aspecte.	Optimistă
16. *Visibility*	Vizibilitate: Știința datelor oferă o privire pătrunzătoare asupra problemelor complexe legate de date vaste.	
17. *Vivify*	Vivacitate: Știința datelor are potențialul de a anima toate tipurile de procese de luare a deciziilor și de afaceri, de la publicitate până la detectarea fraudelor.	Optimistă
18. *Vocabulary*	Vocabular: Știința datelor oferă un vocabular tipic pentru adresarea unei mari varietăți de probleme. Diferite abordări de modelare tratează diverse domenii ale problemelor, iar diferitele tehnici de validare întăresc aceste abordări în diferite aplicații.	Optimistă
19. *Voice*	Voce: Știința datelor oferă posibilitatea de a vorbi în cunoștință de cauză pe o gamă diversă de subiecte.	Optimistă
20. *Voyage*	Voiaj: Să continuăm să învățăm în timp ce abordăm problemele pe care le oferă știința datelor.	Optimistă
21. *Veracity*	Veracitate: Reproductibilitatea este esențială pentru o analiză precisă.	
22. *Volatility*	Volatilitate: Riscul de volatilitate al datelor există, mai ales în sistemele de producție, unde pot apărea situații privind dispariția bruscă a unor date sau formatarea lor necorespunzătoare – situații care trebuie preîntâmpinate.	Pesimistă (erori)
23. *Varmint*	Parazit: Pe măsură ce Big Data devin tot mai vaste, apare riscul ca programele software să conțină erori.	Pesimistă (erori)
24. *Vagueness*	Neclaritate: Sensul datelor rămâne adesea foarte neclar, indiferent de cât de multe date sunt disponibile.	Pesimistă (neînţelegere)
25. *Viability*	Viabilitate: Este dificil să construiești modele robuste și este și mai greu să construiești sisteme care să fie viabile în producție.	Pesimistă (irelevanţă)
26. *Variability*	Variabilitate: Știința datelor modelează adesea surse variabile de date. Modelele implementate în producție pot întâlni date atipice.	Neutră
27. *Variety*	-Varietate: În știința datelor, lucrăm cu multe formate de date (fișiere, baze de date relaționale, rețele de grafice) și cu diferite niveluri de completitudine a datelor.	Neutră
28. *Varnish*	Poleire: Modul în care utilizatorii finali interacționează cu munca	Neutră

	noastră contează, precum și poleiala datelor.	
29. *Vastness*	Vastitate: Odată cu apariția internetului lucrurilor, amploarea datelor mari crește accelerat.	Neutră
30. *Vault*	Seif: Odată cu multitudinea aplicațiilor de știință a datelor bazate pe seturi de date mari și adesea sensibile, securitatea datelor este din ce în ce mai importantă.	Neutră
31. *Velocity*	Viteză: Nu numai că volumul de date crește mereu, dar și rata de generare a datelor (legate de internetul lucrurilor, rețelele de socializare etc.) este în creștere.	Neutră
32. *Venue*	Locul de desfășurare: Munca în domeniul științei datelor se desfășoară în diferite locuri, cu diferite aranjamente: local, în cadrul stațiilor de lucru ale clienților, și în cloud.	Neutră
33. *Verdict*	Verdict: Pe măsură ce un număr tot mai mare de oameni sunt afectați de deciziile modelelor, Veracitatea și Validitatea devin din ce în ce mai importante.	Neutră
34. *Version*	Versionare (controlul versiunilor succesive ale datelor și modelelor): Și tu faci controlul versiunilor, nu-i așa?	Neutră
35. *Victual*	Îndestulător: Big data – hrana care alimentează știința datelor.	Neutră
36. *Viral*	Viral: Cum se răspândesc datele printre alți utilizatori și aplicații?	Neutră
37. *Viscosity*	Vâscozitate: Legată de viteză – cât de dificil este să lucrezi cu datele?	Neutră
38. *Visualization*	Vizualizare: Adesea, este singurul mod în care clienții interacționează cu modelele.	Neutră
39. *Vogue*	Modă: Învățarea automată care devine inteligență artificială , care devine...?	Neutră
40. *Volume*	Volum: Tot mai multe persoane folosesc dispozitive de colectare a datelor, pe măsură ce tot mai multe aparate devin smart. Volumul datelor crește într-un ritm uluitor.	Neutră
41. *Voodoo*	Voodoo: Știința datelor și Big Data nu sunt voodoo, dar cum putem convinge clienții potențiali de valoarea științei datelor, pentru a oferi rezultate cu impact în lumea reală?	Neutră
42. *Vulpine*	Viclenie: „Nate Silver ar vrea să fii o vulpe, te rog.“	Neutră

Există o dezbatere consistentă privind valoarea științei datelor (*data science*). Controversa devine tot mai intensă pe măsură ce argumentele se acumulează atât în favoarea, cât și împotriva atribuirilor de valoare cognitivă modelelor bazate pe date vaste. Până de curând, analiza statistică a datelor era orientată spre confirmarea sau infirmarea unor ipoteze explicative bazate pe teorii specifice unui domeniu, precum medicina,

fizica, meteorologia, psihologia, economia sau sociologia. Datele erau atent colecționate și verificate pentru a elimina erorile și a asigura validitatea măsurărilor și reprezentativitatea eșantioanelor. Timp de secole, metodologia științifică a evoluat în așa fel încât a ajuns să garanteze un control cât mai fin al variabilelor și o observație cât mai precisă a fenomenelor studiate. Modelele de analiză erau aplicate urmărind un control cât mai ridicat al specialiștilor asupra transformărilor aplicate informației. Prin compararea minuțioasă a datelor obținute cu ipotezele generate prin teoriile acceptate, specialiștii dobândeau o înțelegere tot mai profundă asupra lumii. Acest stil de cunoaștere este radical schimbat în știința datelor. După cum am văzut, datele vaste scapă metodelor tradiționale de control, prin volumul și viteza lor de acumulare, precum și prin lipsa de structură. Adesea, modelele de analiză nu mai pornesc de la ipoteze întemeiate teoretic în diferite discipline științifice, ci se ancorează doar în tiparele, reale sau iluzorii, găsite în date. Problema reprezentativității datelor (sau a lipsei acesteia) este mascată de volumul imens de informații. Apare riscul unei confuzii între date vaste și date totale, atunci când credem că informațiile din rețelele de socializare sau motoarele de căutare sunt reprezentative pentru întreaga populație. Valoarea explicației cauzale este înlocuită de valoarea coincidenței și corelației.[8] Totuși, după cum observă Danah Boyd și Kate Crawford[9], două cercetătoare specializate în știința datelor, persistă o mitologie și o idealizare a acestor tehnologii, mitologie care ascunde limitările, aproximările și compromisurile necesare pentru a le face să funcționeze.

În umbra succeselor predictive ale datelor vaste și inteligenței artificiale se află discuțiile privind limitele acestui tip de modelare a lumii și erorile sale caracteristice. În cele ce urmează, vom discuta criticile care denunță sursele si rezultatele dăunătoare ale acestui deficit de cunoaștere. Putem identifica mai multe forme de eroare specifice Big Data:

– *deficitul de precizie*, sau erorile și *bug*-urile inerente unui volum amplu de date, precum și unui volum enorm de linii de cod necesare procesării acestor date;

– *deficitul de înțelegere*, care derivă atât din natura masivă și eterogenă a datelor, cât și din soluțiile de învățare automată (*machine learning*) folosite

pentru a le procesa;

– *deficitul de înțelepciune*, rezultat din reducționismul inerent politicilor bazate pe Big Data;

–*deficitul de adevăr factual* amplificat prin produse și servicii bazate pe Big Data, de la *deep fakes* la *fake news* propagate în rețelele digitale care sunt susținute de capitalismul datelor.

În cele ce urmează voi ilustra fiecare dintre aceste deficite și voi discuta implicațiile lor etice.

Deficitul de precizie: erori ale cunoașterii bazate pe Big Data

Dacă pui o lingură de vin într-un butoi de apă de canal,
obții apă de canal.

Dacă pui o lingură de apă de canal într-un butoi de vin,
obții apă de canal.

(Legea lui Schopenhauer privind entropia,
conform legilor lui Murphy)

Riscul de a include erori factuale derivă din mai multe trăsături specifice ale Big Data. În primul rând, prin chiar *volumul lor masiv*, aceste date nu pot fi verificate la un nivel ridicat de detaliu. Mai mult, erorile tind să se acumuleze, deoarece *Big Data sunt aditive și au o persistență ridicată a informațiilor*.

În al doilea rând, *Big Data acumulează informații din surse eterogene*, al căror conținut este de calitate diferită. Aceasta este cu atât mai mult o problemă a datelor agregate prin brokerii de date, care combină fluxuri online din aplicații și platforme diverse cu surse offline. Pot apărea atât probleme la agregare (datele unui individ să fie combinate cu ale unui alt individ), cât și probleme legate de calitatea slabă a unora dintre datele de pornire, care ajung să polueze fluxurile cu calitate mai ridicată.

Cât de grave sunt însă aceste erori? În statistică, erorile aleatorii sunt diferențiate de erorile sistematice prin faptul că primele nu produc

distorsiuni în concluziile obținute, ci doar imprecizie. Erorile sistematice, precum sub- sau supra-reprezentarea unor valori față de celelalte, pot afecta grav validitatea rezultatelor obținute. Big Data sunt acuzate de ambele tipuri de distorsiuni.

Erorile sunt agravate de precaritatea buclelor de feedback negativ, de autocorectare, alături de buclele de feedback pozitiv, prin care impactul se transmite și se amplifică. Pe de o parte, temporalitatea tot mai accelerată a deciziilor bazate algoritmic înseamnă că anumite decizii greșite se pot propaga mult mai rapid decât pot oamenii percepe și reacționa: „Putem acum face greșeli de calcul catastrofice în nanosecunde, difuzându-le la scală globală. Am pierdut echilibrul dat de timpii de întârziere“, observă analista Marcia Richards Suelzer într-un interviu.[10]

În al doilea rând, oamenilor le este adesea imposibil să detecteze erorile din cauza obscurității volumelor enorme de date și a procesării tip cutie neagră specifică soluțiilor de învățare automată și inteligență artificială. Datele de antrenare pentru produsele ML/AI pot fi atât incorecte, incluzând informații false, cât și părtinitoare, supra- sau sub-reprezentând anumite categorii sociale, ceea ce oferă o imagine nereprezentativă a populației în ansamblu.[11] De exemplu, dacă un set de date este construit mai ales pe persoane tinere, dar este folosit pentru a antrena un model care va fi aplicat și persoanelor vârstnice, acestea din urmă vor suferi de pe urma riscului unor clasificări sau diagnoze eronate.

Imprecizia este amplificată și prin deciziile corelaționale luate de produsele bazate pe ML/AI. Acestea sunt antrenate să detecteze diverse tipologii. Chiar și în eventualitatea în care datele de antrenament sunt un eșantion corect și reprezentativ al datelor populației-țintă, anumiți indivizi pot fi atipici, înde–părtați de tendințele centrale ale populației respective. Aceștia pot fi clasificați și judecați eronat. De asemenea, chiar dacă datele sunt corecte, corelațiile pot fi aparente, adică să nu reflecte procese reale de cauzalitate între cele două variabile, ci (co)incidențe specifice unui context social. Astfel de corelații iluzorii pot fi ridicate la rangul de criteriu decisiv pentru viața indivizilor. Mai mult despre acest tip de vulnerabilitate voi discuta mai jos, în secțiunea referitoare la cauzalitate și corelație.

Erorile aleatorii

Măsoară de multe ori și croiește o dată.

Erorile aleatorii introduc zgomot în date şi pot face dificilă identificarea factorilor asociaţi unui fenomen de interes. Mai mult decât atât, în cazul datelor vaste care sunt folosite inclusiv pentru predicţii şi decizii privind indivizii umani, cineva poate fi direct afectat negativ de procesarea automată a profilului său digital. Această situaţie apare atunci când profilul digital include erori, chiar dacă acestea sunt aleatorii, la nivelul de ansamblu al bazei de date.

Un exemplu de astfel de eroare este oferit de cazul unui medic ginecolog care, prin urmele digitale pe care le lasă căutând simptome și soluții ale unor boli, se comportă similar cu cei care suferă de aceste boli: „Ce se întâmplă dacă algoritmii decid că știu ceva despre tine care nu există cu adevărat? Ar putea să te irite! O prietenă a soției mele, care, întâmplător, e doctor, e învinuită zilnic de motoarele de căutare care îi afișează reclame pentru a-și trata bolile venerice inexistente“[12]. Prin urmare, medicul este diagnosticat în profilul său digital ca fiind suferind de aceste afecţiuni.

Dincolo de frustrarea cotidiană a afișării mesajelor unor reclame irelevante, o astfel de eroare poate conduce însă la consecințe grave privind individul. O eroare aleatorie, strecurată în profilul digital, poate fi interpretată ca un semnal negativ de către deciziile algoritmice ale unor ofertanţi de asigurări de sănătate sau ale unor firme ori departamente de resurse umane care procesează cereri de angajare. Individul poate suferi consecințe negative, precum costuri mai ridicate la asigurările medicale, respingerea automată a CV-ului sau scăderea încrederii în sine, fără a putea afla care a fost cauza reală a acestor decizii algoritmice și fără a avea șansa remedierii erorii.

Exemple similare de erori factuale, care pot avea consecințe grave asupra șanselor de angajare, a evaluării credibilităţii financiare sau a riscurilor de sănătate, sunt oferite de jurnalista Frederike Kaltheuner de la Privacy International, în investigația realizată privind propriul profil digital: „Experian UK mă clasifică în funcţie de presupusa mea situaţie financiară (dintr-un motiv inexplicabil, sunt clasificată la categoria «Prosperitate

urbană: bogăție de clasă mondială», firma Acxiom mă plasează într-o categorie numită «Cheltuieli masive pentru alcool acasă» (să fie oare din cauza cumpărăturilor pe care le-am făcut pentru o petrecere onomastică, la mine acasă?) și o companie numită Affinity Answers crede că am o afinitate cu profilul consumatorilor de «scutece și lavete» (greșit, foarte greșit)“[13]. Fiecare dintre aceste clasificări influențează invizibil și impredictibil deciziile algoritmice luate pe baza profilului digital. Clasificarea în categoria „alcoolism“ sau de „tânără mamă“ poate afecta negativ șansele de angajare, prețul asigurării de sănătate, dar și al altor asigurări, precum cea auto sau cea de viață.

Erorile din Big Data pot avea impact nu doar asupra indivizilor, ci și a ofertanților de produse și servicii. Aceștia din urmă pot fi afectați în măsura în care predicțiile eronate conduc la degradarea reputației lor și a încrederii consumatorilor. Un studiu empiric al calității țintirii audiențelor prin date disponibile de la 19 firme specializate în agregarea de informație arată o variabilitate foarte ridicată a preciziei și o frecvență ridicată a erorilor, estimând că raportul cost–beneficiu devine avantajos doar pentru companiile care operează cu bugete ridicate.[14] În raportul Deloitte, „Predictably inaccurate: the prevalence and perils of bad big data“[15], autorii observă că datele ce includ erori pot conduce la probleme reale pentru organizații, mergând de la mici incidente jenante până la alienarea treptată a clienților. Problema apare atunci când, pe baza unor date care conțin informații factuale false, companiile realizează o microsegmentare și o microțintire a clienților, utilizând aceste clasificări potențial eronate pentru decizii de marketing. Mai mult, observă autorii raportului, încrederea aproape oarbă în date poate conduce la renunțarea la precauții și spirit critic privind rezultatele analizelor, ceea ce poate fi contraproductiv. Autorii prezintă un studiu privind percepția indivizilor asupra corectitudinii profilurilor lor digitale obținute de la un „broker de date cu poziție de vârf în piață“ (nespecificat), studiu ale cărui concluzii sunt îngrijorătoare. Din cei 107 participanți, peste două treimi evaluează corectitudinea profilurilor lor în ansamblu ca fiind mai scăzută de 50%. Chiar și datele socio-demografice de bază, sau datele privind tipul autovehiculului pe care îl conduc au avut niveluri evaluate de corectitudine foarte scăzute. Așadar,

erorile factuale din Big Data nu sunt doar evenimente izolate, ci sunt o problemă de scară largă a domeniului.

O problemă recurentă identifică și studiile pe bazele de date folosite în antrenarea algoritmilor de inteligență artificială. De exemplu, un studiu recent ajunge la concluzia că bazele de date folosite pentru testarea produselor de ML/AI sunt foarte adesea afectate de erori extinse. Conform estimării acestui studiu, există aproximativ 3,4% intrări eronate, în medie, pentru cele 10 baze de date vaste studiate; mai mult, baza de date de validare derivată din ImageNet, folosită frecvent pentru soluțiile de recunoaștere a imaginilor, are 6% intrări eronate.[16] Sursa erorilor provine, în mare parte, din faptul că datele sunt etichetate de o forță de muncă vastă, anonimă și slab plătită, prin platforme de crowd-working.[17]

Abordarea legală menită să răspundă problemei erorilor aleatorii din profilurile digitale personale poate fi găsită în GDPR.[18] Directiva europeană stipulează mai multe drepturi ale persoanelor relevante în acest sens.[12] În primul rând, este vorba despre dreptul de a afla dacă o organizație procesează datele personale și dreptul de a obține aceste date (vezi art. 15, Dreptul la acces al subiectului datelor). Apoi, indivizii au dreptul la rectificarea neîntârziată a datelor personale imprecise sau incomplete (vezi art. 16, Dreptul la rectificare). Nu în ultimul rând, este stipulat dreptul la restricționarea procesării datelor personale atunci când corectitudinea acestora este disputată (vezi art. 18, Dreptul la restricționarea procesării).

Aceste drepturi pot fi folosite de indivizi, precum și de reprezentanții lor, pentru a verifica și pentru a corecta datele factuale incorecte pe baza cărora organizațiile private sau publice îi clasifică și îi judecă. Desigur, cetățenilor le este practic dificil să afle situațiile în care datele despre ei au fost afectate de erori. Puțini dintre indivizi au resursele de timp și competențele necesare pentru a-și cere și a-și verifica datele colectate de companiile specializate în agregarea informațiilor. Totuși, aceste drepturi asigură posibilitatea de verificare și de remediere atunci când se ivesc situații care par, la prima vedere, problematice, precum și o metodă de a stimula companiile și organizațiile să ia măsuri preventive în vederea asigurării corectitudinii datelor cu care operează.

Erorile sistematice

Judecătorul când îți e potrivnic, cu cine o să te judeci?

Într-o după amiază de joi, ianuarie 2020, Robert Julian-Borchak Williams a primit un apel telefonic la serviciu. Conform relatării din New York Times, bazată pe investigații ulterioare și pe implicarea Uniunii Americane pentru Libertăți Civile[19], o persoană care se pretindea ofițer al Poliției Detroit i-a solicitat să vină la secție pentru a fi arestat. Williams a crezut că e o glumă proastă și, desigur, și-a văzut de slujbă și de viață. O oră mai târziu, când a ajuns acasă, o mașină de poliție s-a oprit în fața locuinței sale și doi ofițeri l-au oprit, i-au pus cătușe și l-au arestat, în fața soției și a fiicelor lui. Nu i-au fost oferite prea multe detalii. Când soția a întrebat de ce îl arestează, unul dintre ofițeri i-a răspuns să caute pe Google. Williams a petrecut noaptea în detenție, dormind pe podeaua unei celule supraaglomerate, lângă coșul de gunoi. Vineri, la prânz, a fost dus într-o cameră de interogare, unde a aflat, pentru prima dată, ce acuzație i se aduce. Doi detectivi l-au întrebat dacă el e cel ce a jefuit un magazin Shinola, un butic de lux care vinde ceasuri, biciclete și genți de piele. Detectivii i-au arătat două fotografii preluate de o cameră de luat vederi, dintre care una arăta un bărbat cu o șapcă în timp ce fura din magazin cinci ceasuri, în valoare de 3.800 de dolari. A doua fotografie prinsese chiar chipul bărbatului, puțin încețoșat, dar de la apropiere. Deși bărbatul era negru, ca și Williams, era destul de evident că e o altă persoană. „Aici ești tu?“ au întrebat polițiștii. „Nu, nu sunt eu“, a răspuns Williams. „Credeți că toți bărbații negri arată la fel?“ Acesta a fost, probabil, primul caz de arest nejustificat din SUA bazat pe dovezi generate de un algoritm de recunoaștere facială. Poza hoțului a fost transmisă poliției, care a introdus-o într-o bază de date, unde algoritmul a identificat mai mulți candidați posibili pornind de la arhiva fotografiilor permiselor de conducere. Williams a fost unul dintre ei. Fără alte investigații suplimentare, care ar fi dovedit că Williams avea un alibi pentru momentul furtului, poliția a arătat pozele unei angajate a buticului, care l-a identificat pe Williams ca fiind hoțul. Ca urmare a acestui incident, Williams a petrecut 30 de ore în arest și a trebuit să plătească 1.000 de dolari cauțiune pentru a fi eliberat, urmând apoi să-și caute un avocat care să îl reprezinte,

contra unui cost de aproximativ 7.000 de dolari. Fiind însă un caz emblematic pentru riscurile recunoașterii faciale, Uniunea Americană pentru Libertăți Civile i-a preluat dosarul, pentru a-l reprezenta în justiție. Greșeala petrecută în cazul lui Williams nu a fost însă aleatorie. În spatele acestui incident se află o variabilă ce explică parțial atribuirea eronată a delictului: anume culoarea pielii.

Spre deosebire de erorile aleatorii, care introduc zgomot în date fără a afecta tiparele observate, *erorile sistematice pot conduce la detectarea unor tipare iluzorii*. Astfel de erori sistematice sunt una dintre cauzele discriminării algoritmice. Acest tip de discriminare apare în decizii care se bazează, parțial sau total, pe procesarea automată a profilurilor digitale.

De exemplu, numeroase studii au documentat faptul că *algoritmii de recunoaștere facială utilizați până în prezent sunt, prin deciziile pe care le iau, sexiști și rasiști*. Mai precis, acești algoritmi sunt pasibili de erori mai ridicate în recunoașterea fețelor femeilor, față de cele ale bărbaților și a persoanelor de culoare, față de persoanele albe. Acest fenomen se petrece deoarece bazele de date pe care algoritmii de recunoaștere facială au fost antrenați conțin semnificativ mai multe fețe masculine albe decât de alte tipuri.[20] Acestea sunt exemple de erori sistematice, afectând disproporționat anumite categorii sociale.

Dat fiind că algoritmii de recunoaștere facială sunt utilizați în SUA, inclusiv de către poliție, pentru identificarea potrivirilor între fețele suspecților și fotografiile furnizate de bazele de date ale poliției sau de camerele de supraveghere instalate în spațiul public, aceste erori pot conduce la identificarea și acuzarea eronată a unor persoane. Probabilitatea de imputare pe nedrept este, așadar, mult mai mare pentru persoanele de culoare – ceea ce documentează și raportul „The perpetual line-up“. În 2018, Uniunea Americană pentru Libertăți Civile a scanat fețele membrilor Congresului SUA cu soluția de recunoaștere facială a Amazon, Rekognition. Rezultatul a constat în identificarea, printre membrii Congresului SUA, a 28 de false potriviri pozitive cu suspecți de infracțiuni, reprezentând, totodată, disproporționat, persoane de culoare.[21] Identificăm aici atât un nivel ridicat al erorilor, cât și părtinirea rasială a acestora.

Ceea ce pune sare pe rană este faptul că *erorile nu au rămas la stadiul de*

studiu științific, ci provoacă daune reale indivizilor. De exemplu, până în 2021, sunt cel puțin trei cazuri documentate de bărbați afro-americani care au fost arestați în mod eronat în SUA, pe baza indicațiilor oferite de programe de recunoaștere facială.[22] Aceste persoane au suferit daune reale, inclusiv financiare, emoționale și morale. Există însă multe alte arestări bazate pe potriviri realizate de astfel de programe, unele dintre ele conducând la condamnarea la închisoare a indivizilor. Câte dintre aceste condamnări, în care recunoașterea facială joacă un rol decisiv, sunt corecte și câte sunt eronate? Uniunea Americană pentru Libertăți Civile estimează că, la o rată de eroare de unu la mie, ar fi posibil ca, doar în Florida, să fie arestate eronat aproximativ opt persoane în fiecare lună.[23] Aceste persoane ar fi, foarte probabil, afro-americani, deoarece, pe de o parte, programele au rate de precizie mai slabe pentru fețele care nu sunt albe și, pe de altă parte, poliția americană procesează și arhivează proporțional mult mai multe fotografii ale persoanelor de culoare.

Folosind tehnologii similare, guvernul chinez utilizează soluții de recunoaștere facială și de detectare a etnicității pentru supravegherea și controlarea continuă a minorității uigure din provincia Xinjiang, așa cum am menționat și în capitolul 3. Prin implementarea acestor metode de control, cetățenii uiguri sunt supuși la încălcări grave ale drepturilor omului, de la detenție în lagăre de reeducare la încălcarea persistentă a libertăților umane fundamentale.[24]

Erorile aleatorii și sistematice din datele de antrenament și din deciziile luate de algoritmii de recunoaștere facială au condus la sporirea neîncrederii publice față de utilizarea lor în general și, mai ales, în cadrul poliției. De exemplu, IBM a anunțat în 2020 că se retrage de pe piața recunoașterii faciale, iar Amazon și Microsoft au decis să nu mai permită utilizarea soluțiilor lor de către forțele responsabile cu menținerea ordinii publice.[25]

În iunie 2020, comitetul de politică din domeniul tehnologiei a prestigioasei asociații profesionale inginerești din SUA, ACM (Association for Computing Machinery), a declarat public: „[Comitetul nostru] solicită cu insistență suspendarea imediată a utilizării prezente și viitoare, private și publice, a tehnologiilor de recunoaștere facială în toate circumstanțele cunoscute sau rezonabil predictibile care conduc la prejudicii privind

drepturile umane şi legale stabilite“[26]. Totuşi, alte organizaţii continuă să dezvolte şi să ofere soluţii comerciale de recunoaştere facială autorităţilor sau companiilor interesate să le utilizeze.

Ca şi alte tehnologii controversate, de la detectorul de minciuni la medicina alternativă, recunoaşterea facială extinde mult puterea celor care dispun de această metodă, făcând din ea o resursă foarte atractivă, în ciuda daunelor colaterale pe care le produce în privinţa drepturilor omului. Este necesară o mobilizare socială amplă, cu susţinere legală, dacă dorim să evităm generalizarea unei astfel de tehnologii de supraveghere.

Deficitul de înțelegere: iluzia cunoașterii

Un nebun aruncă o piatră în baltă, şi zece cuminţi n-o pot scoate.

Chiar dacă facem abstracţie de erorile din date şi din programele care le procesează, problema deficitului de cunoaştere persistă. Deficitul de cunoaştere, sub forma deficitului de înţelegere, reprezintă o formă mai gravă decât cele discutate până acum. Acesta apare la mai multe niveluri.

Spre deosebire de bazele de date analizate statistic, ce permit un anumit nivel de explorare umană, *volumele de date acumulate şi procesate sunt atât de mari şi de eterogene, încât nu pot fi cuprinse cu mintea*. Datorită vastităţii cantitative, dar şi datorită lipsei unei structuri şi a caracterului dinamic presupus de acumularea continuă, datele vaste ies din tiparele analizelor statistice clasice. Big Data pot fi prelucrate şi interpretate doar în echipe transdisciplinare, intens mediate tehnologic, apelând la metode de analiză şi, mai ales, de vizualizare la scară largă, prin care datele devin inteligibile uman.[27]

Volumele de date sunt în continuă creştere, iar algoritmii înşişi sunt construiţi pe baza altor algoritmi. Logica generală este una a rezolvării problemelor complexe prin sporirea complexităţii soluţiei, în detrimentul inteligibilităţii. Cercetătorii şi experţii din domeniu îşi utilizează competenţele pentru a construi metode tot mai sofisticate de analiză şi de vizualizare, cu obiectivul de a valorifica date tot mai vaste. Imperativul cantităţii se traduce într-un imperativ al complexităţii, care conduce la sacrificarea inteligibilităţii umane: „Algoritmul nu mai poate fi complet

înţeles – sau cunoscut – de o singură persoană, tocmai din cauză că a devenit un mon-stru al datelor care necesită o hrănire continuă“[28].

Acest deficit de înțelegere este mascat și amplificat de iluzia obiectivității algoritmilor. La un nivel ideal de abordare, algoritmii conotează o certitudine matematică, datele vaste conotează o acoperire a fenomenului, iar intersecţia celor două comunică simultan obiectivitate şi cunoaştere comprehensivă. Totuși, algoritmii nu sunt altceva decât „opinii inserate în cod“, observă O’Neil în discursul său intitulat „Era încrederii oarbe în Big Data trebuie să înceteze“[29]. Tot O’Neil observă că această încredere în veridicitatea concluziilor derivate din Big Data este parţial întemeiată pe gândirea mitică, delegând responsabilitatea pentru decizie unei autorități iluzorii a datelor: „Putem defini Big Data ca o combinație între trei lucruri: tehnologie, analiză și mit [...]: oamenii au un fals sentiment de încredere în date, mai ales pe măsură ce bazele de date devin tot mai mari și mai mari“[30].

Principala acuzație adusă Big Data și deciziilor algoritmice constă, prin urmare, nu doar în riscul generării de probleme, precum clasificări eronate sau discriminatorii ale indivizilor, ci și în aparenta imunitate a acestui proces la feedbackul negativ. Știința datelor și automatizarea prin ML/AI este însoțită de o veritabilă mitologie a valorii științifice și predictive a acestor tehnologii, care riscă să orbească organizațiile statale, comerciale sau politice privind neajunsurile și limitările datelor.

Exacerbând viziunea pozitivistă, în care cunoașterea precisă este posibilă inclusiv în studiul societăţilor prin analiza matematică a datelor, știința datelor induce un exces de certitudine în situații în care incertitudinea și deschiderea spre viitoruri multiple ar putea fi o soluție mai bună, ducând la un echilibru moral mai stabil.

Cauzalitate și corelație

Surdul n-aude, dar le potrivește.

În secolele trecute, exista obiceiul de a preveni răceala purtând nişte căţei de usturoi atârnaţi la gât. Și astăzi usturoiului îi sunt atribuite numeroase virtuţi protective, dar cu condiția de a-l mânca. La ce ar putea folosi purtarea sa pe post de medalion? Obiceiul făcea parte din tradiția utilizării

medicale a plantelor aromatice, analog cu masca cu cioc a medicilor europeni în perioadele de ciumă. Se presupunea că aromele puternice (fie ale usturoiului, fie ale substanţelor din ciocul măştii) purifică aerul de miasmele purtătoare de boli. Desigur, ştim azi că răceala e provocată de virusuri, nu de miasme.

Să fi fost aceasta doar una dintre nenumăratele superstiţii prin care oamenii simţeau că îşi controlează viaţa, fără vreo eficacitate reală? Este, desigur, plauzibil. Totuşi, putem identifica azi şi un mecanism cauzal real, dar diferit de cel presupus în trecut. Oamenii care purtau usturoi la gât miroseau ei înşişi neplăcut şi îi ţineau la distanţă pe cei din jur, scăzând riscul propriei infectări. Prin urmare, este posibil ca această recomandare medicală străveche să fi fost, de fapt, justificată – dar printr-un cu totul alt temei. În spatele corelaţiei percepute exista o altă cauzalitate decât cea presupusă, ceea ce nu diminua eficacitatea remediului.

Adesea, Big Data permit descoperirea unor corelaţii care sunt folosite pentru a prezice şi controla viitorul, fără a înţelege, de fapt, pe ce se bazează respectivele corelaţii. Spre deosebire de ştiinţa secolelor trecute, care măcar căuta o explicaţie, ştiinţa datelor nu mai insistă pe înţelegerea acestor corelaţii. Sunt suficiente găsirea şi folosirea lor.

Recomandările de acţiune (*insights*) bazate pe Big Data şi *machine learning* reflectă asocieri ale variabilelor, asocieri care nu sunt neapărat validate de o teorie ştiinţifică. *Corelaţia nu este sinonimă cu o relaţie de cauzalitate, dar acest lucru a fost de mult timp acceptat ca un fapt, ba chiar ca o virtute în ştiinţa datelor*.

Presupunerea implicită din epistemologia Big Data este că, dacă găsim predicţii suficient de precise pe baza unor corelaţii, nu mai contează dacă aceste corelaţii ascund factori cauzali neidentificaţi. Tot ceea ce contează este validitatea predicţiei finale, care stă la baza ţintirii reclamelor, preţurilor, mesajelor sau verdictelor personalizate.

Valorizarea predicţiei ca indicator de succes al unui model analitic nu este specifică Big Data, ci provine din epistemologia bazată pe legi cauzale generale şi deducţie, tipică ştiinţelor naturii şi aplicată adesea şi în cazul acţiunilor umane.[31] Desigur, ştiinţele socio-umane au dezvoltat şi epistemologii şi metodologii alternative, care ţin cont de intenţionalitatea

umană, de liberul-arbitru și de semnificațiile culturale distincte ale unui obiect sau fenomen, observă Cathy O'Neil. Știința datelor însă nu a mers în această direcție a înțelegerii acțiunii și semnificațiilor sale, ci, dimpotrivă, a dus la extrem modelul predicției bazat pe deducții din legi cauzale, înlocuind însă legile cauzale prin legi corelaționale. Cu alte cuvinte, observă matematiciana, știința datelor a legitimat o preocupare exclusivă cu aparențele, urmărind succesul predictiv, indiferent dacă înțelegem sau nu cum și de ce se produce evenimentul prezis.

Cutiile negre din deciziile algoritmice, din învățarea automată și din inteligența artificială

În oala acoperită nimenea nu știe ce fierbe.

Confundarea cauzalității cu corelația este agravată de imposibilitatea examinării în detaliu a modelelor predictive care leagă datele de decizii, pentru a înțelege întemeierea rezultatelor. Această imposibilitate derivă din două considerente. Pe de o parte, algoritmii utilizați de organizațiile comerciale, dar și publice, sunt adesea apărați de drepturi de proprietate intelectuală și nu pot fi inspectați.[32] Pe de altă parte, cea mai mare dificultate provine din specificul soluțiilor de învățare automată și inteligență artificială.

În stilul clasic, soluțiile de tip statistic și procedural aveau modelele (oricât de complicate și de mari ar fi fost) scrise de oameni și erau inteligibile citirii umane. Spre deosebire de acestea, soluțiile bazate pe învățare automată (*machine learning*, ML) și inteligență artificială (*artificial intelligence*, *AI*) deleagă sistemelor digitale chiar operațiile de construire a modelului. Oamenii sunt cei ce oferă datele și condițiile de succes; soluțiile ML/AI sunt cele care învață, adaptându-se continuu la feedbackul primit și optimizând metricile care definesc reușita. Așadar, *modelul final la care ajunge o soluție ML/AI nu mai este inteligibil pentru oamenii ce au gestionat întregul proces*.

Chiar și în absența constrângerilor de proprietate intelectuală, modelul este esențialmente o cutie neagră. Știm ce intră în această cutie, putem testa ce iese din această cutie, dar operațiile însele, prin care inputul este conectat

cu outputul, rămân misterioase. Aceasta înseamnă şi că, atunci când o decizie este considerată greşită, este foarte greu, dacă nu chiar imposibil, să fie detectate cauzele sau sursele erorii: deciziile ML/AI nu pot fi supuse unei inginerii inverse sau reconstituite. În ultimii ani, tot mai multă atenţie este acordată acestei probleme, prin cercetări care încearcă să dezvolte inteligenţă artificială explicabilă uman (*explainable AI*).

Un caz care ilustrează această dificultate s-a petrecut în noiembrie 2019, când Apple a lansat un card de credit (Apple Card) în colaborare cu Goldman Sachs. Imediat după lansare, două personalităţi ale lumii IT, David Heinemeier (creatorul Ruby on Rails) şi Steve Wozniak (cofondatorul Apple) au acuzat algoritmii cardului de sexism. Această acuzaţie se baza pe faptul că partenerele lor de viaţă – împreună cu care ei aveau conturi bancare comune – au primit o limită de credit semnificativ mai scăzută decât a lor.[33]

Scandalul public, apărut după postările pe Twitter, a condus la demararea unei investigaţii. La acel moment, observă W. Knight în revista *Wired*, problema nu a constat numai în suspiciunea de părtinire, ci şi în răspunsurile confuze ale celor două organizaţii, care indicau imposibilitatea acestora de a clarifica mecanismul prin care limita de credit a fost calculată: „Răspunsul primit de la Apple doar a adăugat confuzie şi suspiciune. Nimeni din partea companiei nu părea capabil să de-scrie cum anume funcţionează algoritmul, şi cu atât mai puţin să-i justifice outputul. Când Goldman Sachs, banca emiţătoare a Apple Card, a insistat că nu există nici o părtinire de gen în algoritm, nu a reuşit să ofere nici o dovadă“[34].

Aproape un an şi jumătate mai târziu, concluziile investigaţiei realizate de Departamentul Serviciilor Financiare din New York au indicat faptul că algoritmii Apple Card nu discriminează pe baza genului.[35] Investigaţia se bazează pe analiza a 400.000 de solicitanţi din New York, identificând faptul că nu există un tipar direct sau indirect de (de)favorizare în funcţie de gen. A rezultat că alţi factori, care ţin de istoricul de credit şi datorii, au fost responsabili de aceste discrepanţe, identificate în cazurile din 2019. Este interesant, prin urmare, cum *o astfel de situaţie nu a putut fi clarificată chiar la momentul apariţiei sale*, mai ales că a venit de la persoane cu vizibilitate foarte ridicată, stârnind un scandal care a afectat credibilitatea

produsului.

Algoritmii nu sunt doar opaci, ci și efemeri, în sensul în care sunt în continuă schimbare. Spre deosebire de un individ uman, despre care gândim în termenii unei identități legale persistente sau a unei identități psihologice și sociale relativ stabile, algoritmii sunt în continuă modificare, de la o clipă la alta. Un algoritm, care funcționa ieri într-un fel, va putea fi adaptat și modificat sau se va adapta singur, în cazul soluțiilor autonome. Astfel el devine dificil de urmărit în raționamentele pe care le face. Și, mai mult decât atât, algoritmul este dificil de justificat în cazurile în care trebuie pus degetul pe rană cu scopul identificării unei dispoziții, a unei surse stabile a erorilor. Această efemeritate împiedică înțelegerea și îngreunează monitorizarea deciziilor bazate pe Big Data și ML/AI, făcând imposibilă atribuirea responsabilității legale sau morale.

Modelul european de reglementare a politicilor algoritmice include o reacție la problema neinteligibilității deciziilor algoritmice, în articolul 22 din GDPR[13], referitor la profilarea și procesarea automată a datelor privind indivizii. Persoanele au dreptul atât să refuze ca situația lor să fie judecată doar prin procesare automată, cât și să solicite clarificări, să conteste decizia și să obțină o decizie judecătorească cu privire la cazul lor, atunci când procesarea automată nu poate fi evitată conform legii.

Deficitul de înțelepciune: dezumanizarea indivizilor ca *data points*

Până nu iei șaua de pe cal, nu-i vezi rana-n spate.

Putem înțelege cel mai bine profunzimea criticilor adresate Big Data, Data Science și soluțiilor asociate ML/AI ca indicând nu doar erori sau lipsă de înțelegere, ci și o lipsă de înțelepciune. Conceptul de înțelepciune practică, sau phronesis, formulat de Aristotel în Etica nicomahică[36], este foarte relevant pentru a înțelege această acuzație.

Phronesis este un concept mult discutat, care, în urma traducerilor și interpretărilor diferite, trimite spre trei dimensiuni distincte: rațiunea sau justificarea, intuirea situației și urmărirea morală a unei vieți bune.[37] Prin urmare, judecata practică, în sensul de *phronesis* (înțelepciune practică), se referă la capacitatea specific umană de a evalua o situație de viață

particulară în mod raţional şi argumentat, ţinând cont de detaliile acelei situaţii şi de intuiţia privind natura sa, având în vedere nu doar interesele personale sau imediate, ci sensul unei vieţi bune.

În 2021, profesorii Andrew Stephen şi Alex Connock, de la Universitatea Oxford, au făcut un experiment. În cadrul unei dezbateri pe tema eticii inteligenţei artificiale, au invitat şi un participant non-uman, mai precis, modelul de inteligenţă artificială Megatron Transformer, dezvoltat de compania Nvidia.[38] Tema propusă dezbaterii era: „Credem că inteligenţa artificială nu va fi niciodată etică“. Răspunsul inteligenţei artificiale a venit în sprijinul acestei teze, argumentând neutralitatea tehnologiei. Totuşi, observă autorii, aspectul cel mai interesant al opiniei Megatron Transformer nu este direcţia pe care a adoptat-o, ci faptul că poate, la fel de bine şi de convingător, să argumenteze şi în direcţia opusă. Participând într-o serie mai lungă de dezbateri, modelul a oferit opinii de orice parte i-au fost solicitate, cu o coerenţă surprinzătoare. Regăsim aici o temă centrală în discuţiile privind etica inteligenţei artificiale. Modelele ML/AI nu se bazează pe o înţelegere a lumii, adică nu sunt propriu-zis *întemeiate*. Ele funcţionează prin identificarea şi reproducerea formelor matematice de ordine pe care le sesizează în torenţii de numere preluaţi din datele vaste. Cele mai ample modele, capabile să scrie autonom texte sau să poarte un dialog verosimil, sunt lipsite de orice fundament al înţelegerii gândirii sau acţiunii omeneşti. Jitendra Malik, specialist în inteligenţă artificială şi profesor la Universitatea din California, Berkeley, consideră că „aceste modele sunt, de fapt, castele suspendate în aer; nu au nici o fundaţie [...]. Limbajul pe care îl produc aceste modele nu este întemeiat, este bazat pe imitaţie, nu există nici o înţelegere reală“[39].

Putem remarca aici deficitul clar de *phronesis* pentru deciziile bazate pe Big Data şi ML/AI. În primul rând, *acestea nu pot fi justificate*, din cauza proprietăţilor de incomprehensibilitate umană ale modelelor predictive şi ale datelor care le informează. În cazul modelelor ML/AI devine, practic, imposibil să înţelegem de ce un anumit caz particular generează o anumită decizie, chiar şi dacă avem acces direct şi nelimitat la model. În cazul unei inspecţii, complexitatea se transformă în opacitate.

În al doilea rând, deciziile automate se bazează pe tipare foarte largi de

date, care sunt trans-situaţionale: aşa-numitele *Big Data insights.* Acestea din urmă constau în recomandări specifice de acţiune menite să atingă un anumit obiectiv; *nu se referă la înțelegerea nuanțată a unei situații*, ci la perceperea de ansamblu a unor tipare din informaţie provenite din foarte multe situaţii disparate. Această insensibilitate la nuanțele situației face foarte problematică utilizarea ML/AI şi pentru automatizarea deciziilor din viaţa de zi cu zi a oamenilor, şi pentru automatizarea ştiinţei, deoarece ambele domenii necesită decizii fine, bazate pe experienţă şi judecată umană a ceea ce este potrivit unui context foarte specific.[40]

Modelele ML/AI sunt adesea create, antrenate şi testate în condiţii artificiale, de laborator, pe date de calitate superioară celor din lumea utilizatorilor reali. Aceasta este una dintre explicaţiile pentru dificultăţile considerabile cu care se confruntă modelele de inteligenţă artificială în domeniul îngrijirii medicale, observă un reportaj *Wired*.[41] Există domenii înguste în care au fost făcute progrese vizibile, aplicate deja în tratamentul pacienţilor, precum diagnoza unor forme de cancer de piele, boli ale ochilor sau ale inimii. În multe alte situaţii însă, datele care pot fi obţinute din sistemul medical sunt prea puţine sau de prea slabă calitate în raport cu necesităţile modelelor performante ML/AI. Proiecte care par sclipitoare în laborator ajung să eşueze în viaţa reală în domenii critice, de exemplu în alocarea resurselor de sănătate[42], sau, schimbând domeniul, în prevenirea criminalităţii[43]. Odată ce încep să prelucreze informaţiile unor oameni reali, modelele traduc inegalităţile sociale încorporate în date în decizii discriminatorii, mascându-şi lipsa de discernământ prin iluzia obiectivităţii şi a preciziei tehnologice.

În al treilea rând şi poate cel mai important, *modelele ML nu au şi nu pot avea o judecată morală a unei vieţi bune.* Ele sunt esenţialmente *descriptive*, identificând corelaţii şi derivând predicţii pe baza a ceea ce s-a întâmplat deja, sintetizat prin modele statistice. Mai mult, ele descriu realitatea doar prin ceea ce poate fi măsurat, lipsindu-le orice înțelegere reală a fenomenelor pe care le modelează. Ceea ce noi numim o tehnologie de detectare a fețelor este, de fapt, „o tehnologie-de-detectare-a-unui-set-de-pixeli-ce-adesea-corelează-bine-cu prezența-fețelor-în-setul-de-date-pe-care-le-ai-colectat“, observă specialistul în robotică Bill Smart.[44] Claritatea

predicţiilor şi deciziilor este pur matematică, derivată din straturi succesive de calcule fără nici o legătură cu certitudinea unei judecăţi corecte cu privire la o situaţie reală. Dan McQuillan, specialist în inteligenţă artificială şi lector la Goldsmiths, Universitatea din Londra, observă că ştiinţa datelor se întemeiază pe filozofia neoplatonică a realităţii superioare a formelor abstracte, în defavoarea conţinuturilor concrete.[45] Modelele ML/AI generează clasificări care nu reflectă ierarhii de valoare morală, ci doar ordonări numerice preluate din măsurările lumii. Mai mult, aceste calcule sunt într-atât de complicate, încât logica lor devine inaccesibilă oamenilor.

Observăm că ştiinţa datelor, ca orice altă ştiinţă ancorată în universul empiric care generează afirmaţii *factuale*, nu poate emite afirmaţii *normative* fără a cădea în eroarea este/trebuie (*is/ought*), identificată de filozoful David Hume. Totuşi, promotorii deciziilor algoritmice consideră că este posibil ca aceste modele să judece, în practică, valoarea unor comportamente sau a unor indivizi. Această judecată se realizează prin inputurile morale explicite sau implicite din programele de antrenare şi din bazele de date oferite.

În lipsa unor criterii de valoare explicite, modelele AI/ML devin conservatoare şi încearcă să reproducă prin decizii un viitor coerent cu tiparele identificate în trecut. Observăm că, deşi este imposibil ca modele ML/AI să delibereze moral, deciziile lor au consecinţe relevante moral, fie că autorii sau proprietarii lor intenţionează acest lucru sau nu.

Acest deficit de înţelepciune poate fi privit şi ca un reducţionism. Din punctul de vedere al proceselor de decizie algoritmică bazate pe Big Data, indivizii sunt transformaţi în nori de puncte într-un spaţiu n-dimensional de atribute. Desigur, orice cuantificare este reducţionistă, dar transformarea unor modele decizionale automate în agenţi autonomi, nesupervizaţi de oameni, transformă mizele unor astfel de simplificări. Ştiinţa datelor şi produsele care procesează date vaste au extins şi au normalizat practica de a privi indivizii ca elemente ale unei mulţimi, în care individualitatea lor se traduce mai degrabă ca o sursă de eroare decât ca o sursă de valoare. Modelele de analiză şi interpretare specifice pentru Big Data au, mai mult decât modelele statistice clasice, o anumită „neglijenţă sau relaxare“[46] în manevrarea datelor, deoarece individualitatea şi relevanţa unui anumit punct

sunt mult mai scăzute. Berendt *et al*. remarcă şi noul vocabular introdus de ştiinţa datelor şi modelele ML/AI, în care conceptele de noduri şi muchii ale reţelelor sau de predictibilitate a acţiunilor diluează semnificaţia individualităţii.

Andrew Smart şi colaboratorii săi discută chiar această problemă în lucrarea lor „Why Reliabilism Is Not Enough: Epistemic and Moral Justification in Machine Learning“[47]. Autorii observă că, dacă ne raportăm doar la problema validităţii epistemice, poate fi considerat legitim să acceptăm adevărul concluziilor obţinute prin procese pe care nu le putem înţelege, dar despre care observăm că sunt de încredere, că lucrează predictibil. De exemplu, examinând sursele cunoaşterii, avem încredere (de regulă) în percepţiile noastre chiar dacă nu înţelegem circuitele neuronale şi senzoriale ale văzului sau auzului. Aceste cutii negre nu invalidează valoarea de adevăr pe care o acordăm percepţiei. Autorii denumesc acest tip de epistemologie *reliabilism*, pe care o putem traduce printr-o *epistemologie a încrederii suficiente*.

Totuşi, observă ei, această întemeiere a valorii epistemice este insuficientă în registrul moral. O judecată morală este întotdeauna întemeiată, adică trebuie să poată fi pusă sub semnul întrebării şi apărată prin argumente. Toate instituţiile care judecă oamenii şi dau verdicte ce le influenţează viaţa – de la curţile de justiţie la procesele de admitere sau recrutare în şcoli şi locuri de muncă, acceptare sau respingere a cererilor de credit sau asigurare etc. – au proceduri detaliate care permit contestarea deciziilor, reexaminarea temeiurilor în care au fost luate şi, eventual, corectarea lor. Desigur, observă autorii, această cerinţă nu se aplică deciziilor luate prin intermediul modelelor ML/AI bazate pe Big Data.

Prin urmare, *nu putem extinde epistemologia încrederii suficiente la rang de teorie morală*. De exemplu, *nu putem delega unui algoritm judecarea și pedepsirea unui om, fără acordarea unor justificări clare și suficiente*. În concluzie, autorii recomandă un nivel de judecată umană care să asigure valoarea morală a deciziilor ML/AI, inserată ca un fel de înveliş moral care să garanteze justificabilitatea rezultatelor.

Cred însă că racordarea modelelor ML/AI la judecata umană rămâne problematică, din raţiunile discutate anterior privind natura de cutie neagră

a acestora. În lipsa unor soluții reale de asigurare a inteligibilității modelelor – prin monitorizarea atentă a rezultatelor lor și prin studiul deciziilor luate pe anumite scenarii tipice și cazuri particulare – *se poate ajunge la situația în care așa-numita justificare umană să fie, de fapt, o raționalizare a unei decizii automate neinteligibile*. Oamenii sunt capabili să justifice decizii pe care nu le-au luat, pe care nu le înțeleg și chiar decizii aparent absurde. Cu atât mai mult, dacă este vorba despre o decizie care nu se abate mult de la ceea ce pare normal – așa cum e cazul atunci când un membru al unei categorii dezavantajate continuă să fie dezavantajat –, justificarea umană poate să ofere o aparență de inteligibilitate, fără să controleze realmente procesul.

Lipsa de înțelepciune a deciziilor automate este amplificată, într-un cerc vicios, la nivel organizațional. Pe măsură ce organizațiile deleagă tot mai multe decizii sistemelor automate și investesc propriul capital și propria imagine în soluții bazate pe Big Data, apare o valorizare mitologică a relevanței și corectitudinii rezultatelor acestora. Incertitudinile și nevoia de corectare a unor greșeli ajung să fie traduse în apeluri către și mai multe date, generând o frenezie colectivă a acumulării de informații și a procesării lor tot mai complexe. Astfel, incertitudinea și ignoranța nu mai sunt trăite, organizațional, ca apeluri la prudență, ci ca apeluri la noi proiecte de analiză de date.[48]

Acest cerc vicios este perpetuat prin activitatea industriei ce oferă servicii și soluții de date, care întreține mitul ignoranței continue ce poate fi ușurată doar prin acumulare și consum de și mai multe date. *Observăm aici o capitalizare a ignoranței, o transformare a acesteia într-o sursă de profitabilitate*. Schwarzkopf observă[49] că, în ciuda discursurilor care preamăresc valoarea datelor, de fapt valoarea comercială reală este derivată din afirmarea unei ignoranțe periculoase, ca poziție care necesită intervenția prin știința datelor. Ca multe alte nevoi, nevoia de date este construită social și, observă autorul, este amplificată profesional de experții care vând servicii de producere și modelare a datelor.[50]

Prin urmare, observăm cum procese de feedback pozitiv, adică de amplificare, transformă reducționismul inerent al analizelor statistice într-o problemă de dezumanizare și lipsă de înțelepciune practică (*phronesis*),

agravată organizațional prin deprinderile de acumulare continuă de date, sporire a complexității modelelor și delegare a deciziilor spre sisteme automate de procesare a Big Data.

Deficitul de problematizare etică a științei datelor

Ce știe popa nu trebuie să știe satul.

Care sunt provocările pe care știința datelor le ridică pentru etica științei? Modelele clasice de formulare a riscurilor etice ale științei și de apărare a drepturilor omului în cadrul activității științifice se dovedesc tot mai inadecvate, odată cu avansul științei datelor și a produselor ML/AI care folosesc aceste date vaste.

După cum voi discuta în rândurile următoare, premisele centrale ale eticii științei, orientate spre modelul experimental din științele biomedicale, nu mai sunt valabile în această nouă paradigmă științifică și tehnologică. De aceea, observăm că apar noi vulnerabilități ale indivizilor participanți în această formă de cunoaștere și acțiune, precum și noi responsabilități ale cercetătorilor și ale celor care organizează programele de cercetare bazate pe date vaste.

Știința datelor este un câmp în continuă dezvoltare atât în mediul academic, cât și în mediile de interes public și privat, precum domeniile militare și de securitate, dar și cele comerciale. Acest tip de știință se referă, pe scurt, la analiza și modelarea unor volume vaste de date sau Big Data, adesea prin tehnologii bazate pe învățare automată. Specificul acestei discipline emergente este conferit de tipul datelor și de tehnologiile de analiză – spre deosebire de disciplinele tradiționale definite de obiectul de studiu și de teoretizarea acestuia. Știința datelor poate modela informații din orice domeniu tematic, de la psihologie și sociologie la climă sau codul genetic, putând, de asemenea, surprinde relații între fenomene tradițional alocate unor discipline diferite. Din acest punct de vedere, *știința datelor este a-teoretică*, ceea ce reprezintă și principala *obiecție epistemologică* la adresa ei.[51]

Specificul științei datelor constă în căutarea de tipare în Big Data, obținând corelații, clasificări, rețele și alte forme de ordonare a punctelor

din norii de date. Vorbim despre fluxuri de date care privesc fenomene dintre cele mai diverse, produse în scop de cercetare sau cu alte obiective, agregate de la furnizori diferiți, cu forme eterogene de stocare, în acumulare continuă. Reiterez faptul că în această lucrare tratez, în special, etica științei datelor: adică situațiile în care datele se referă la oameni.

Spre deosebire de bazele de date tradiționale, Big Data au adesea efecte și proprietăți surprinzătoare. Aceste particularități ale datelor vaste provin din combinații neașteptate ale unor fluxuri diferite de informație, din utilizări în noi contexte ale unor date anterioare sau din descoperirea unor proprietăți sau relații anterior neconștientizate. Acest factor-surpriză specific pentru Big Data joacă un rol important în problematizarea etică a științei datelor, deoarece surprizele pot fi benefice, dar, desigur, ele pot fi și disruptive sau dăunătoare unor indivizi sau grupuri.

Putem înțelege provocările științei datelor pentru etica științei dacă avem în vedere faptul că principalele modele de evaluare și reglementare a științei au fost elaborate pornind de la știința biomedicală și, mai precis, de la experimentele cu subiecți umani.[52] Principala provocare la care încearcă să răspundă codurile instituționale de etică a științei este reprezentată de evaluarea, prevenirea și tratamentul riscurilor asociate apariției unor daune. Am în vedere daunele realizate de cercetători asupra subiecților umani cu care interacționează în timpul cercetării, eventual chiar prin intervenția experimentală directă.

Aceste riscuri sunt amplificate de asimetria cunoașterii de informații care apare între cercetători și subiecții participanți. Aceștia din urmă, cel mai adesea, și mai ales în domeniul biomedical, nu înțeleg natura sau consecințele unor tratamente realizate asupra lor. Mai mult, dacă cercetătorii sunt în același timp și doctori, apare riscul unei confuzii între cele două roluri. În timp ce medicul are, în primul rând, o datorie față de pacient, conform jurământului lui Hipocrate, cercetătorul are, în primul rând, o datorie față de umanitate, conform principiilor eticii științei.

Acest conflict este vizibil atunci când, de exemplu, în cadrul unui experiment de tipul RCT, un medic trebuie să administreze un placebo unui pacient, fiind conștient de acest lucru și crezând, din experiența sa clinică (fără să fie disponibile rezultatele studiului în care participă), că tratamentul

experimental ar putea salva viața pacientului. Astfel de situații apar regulat în practica studiilor medicale și au fost dezbătute pe larg în cazul pandemiei Covid-19, odată cu studiile clinice privind tratamentele experimentale propuse pentru combaterea noului coronavirus, realizate în 2020.[53]

În astfel de studii, problematizarea etică apare mai ales în etapa culegerii datelor – etapa în care subiecții sunt supuși intervențiilor potențial invazive sau riscante. De asemenea, acest model de cercetare implică, în mod tradițional, o separare clară între cercetarea propriu-zisă, care constă în culegerea/generarea datelor, analiza lor și publicarea rezultatelor, și alte intervenții ulterioare cercetării, dar care folosesc rezultatele obținute.[54] În modelul clasic al științei care precedă intervenția, etica științei se ocupă cu precădere de cercetarea propriu-zisă, separat de etica intervențiilor ulterioare, ce devine apanajul deontologiei profesiunilor respective.

După cum observă Metcalf și Crawford[55], studiile din știința datelor diferă radical de acest model, în mai multe privințe:

– *Adesea datele sunt preexistente analizei și nu sunt produse de cercetători*, fiind uneori generate de companii private prin activitatea lor comercială, care nu este considerată activitate de cercetare și nu este, așadar, reglementată. De exemplu, Facebook dispune de cantități vaste de date privind comportamentul utilizatorilor, pe care le poate oferi cercetătorilor spre analiză.

– *Datele sunt adesea disponibile public pe diferite platforme sociale sau pot fi accesate de public în virtutea dreptului la informare.* Studiile din știința datelor, de exemplu, pot folosi postările și acțiunile utilizatorilor din rețelele de socializare, pot folosi urmele digitale ale interacțiunilor cu autoritățile publice sau cu diferite organizații pe platforme publice (de exemplu, comentarii pe paginile unor ziare, organizații etc.), pot apela la date anonime sau anonimizate înregistrate de senzori urbani sau de supraveghere în spații publice etc.

– *Datele sunt anonime sau anonimizate*, adică indivizii descriși în bazele de date nu sunt referiți prin numele lor sau alți identificatori personali.

– *Datele sunt vaste și sunt procesate cantitativ și statistic* în scopul identificării unor tipare, corelații, tipuri etc.

Este important de observat aici că, tradițional vorbind, studiile statistice

pe eșantioane mari de indivizi ale căror date au fost preluate indirect, precum în studiile demografice, sau prin proceduri anonime și de interacțiune scurtă, precum anchetele sociologice și sondajele de opinie, nu au fost considerate la fel de intruzive sau riscante ca studiile experimentale din biomedicină, în care culegerea datelor pune un risc direct asupra participanților. Prelucrarea statistică oferea, de regulă, un strat adițional de protecție subiecților, care nu ar fi putut fi identificați din articolele ce descriau modelele de regresie, de analiză factorială sau de clasificare, în care indivizii deveneau doar cifre după virgulă.

Prin urmare, datorită acestor specificități și datorită asocierii cu un tip de cercetare considerat lipsit de riscuri pentru subiecții umani, studiile din știința datelor au fost insuficient problematizate etic, cel puțin, în primele etape de dezvoltare ale acestui nou câmp de studiu. Exemplele de cercetări în care riscul pentru subiecți sau chiar daunele aduse acestora erau vizibile nu au întârziat să apară, conducând la nevoia unei redefiniri a codurilor de etică a științei.

În acest sens, Metcalf și Crawford dau exemplul unei cercetări[56] în care autorii au folosit date disponibile public pentru a evalua plauzibilitatea unei ipoteze culese din mass-media privind identitatea reală a artistului Banksy. Hauge *et al*. au folosit date geografice disponibile public pentru a urmări în detaliu deplasările unui cetățean în relație cu domiciliul său inferat și în relație cu locul în care apăreau graffitti-urile lui Banksy, ajungând la concluzia că este plauzibil ca acest cetățean să fie însuși artistul.

Justificarea etică pe care autorii o oferă pentru legitimitatea muncii lor de urmărire detaliată a vieții unei persoane private constă în faptul că au folosit date publice, precum și în posibila utilitate a metodei propuse în contextul combaterii terorismului. Totuși, este clar că un astfel de studiu poate aduce prejudicii însemnate, materiale și simbolice, artistului, care a ales anonimitatea ca strategie centrală a imaginii sale profesionale.

Mai mult, cercetarea poate aduce prejudicii și cetățeanului urmărit, care, independent dacă este sau nu Banksy, va intra în atenția publică într-un mod nedorit. Nici una dintre persoanele prezentate în studiu nu și-a dat consimțământul pentru această utilizare a datelor publice disponibile și este puțin probabil că și l-ar fi dat în cazul în care ar fi fost întrebate. Observăm

aici că specificul studiilor pe baza Big Data constă într-un aparent profil inofensiv al datelor și al metodei, care se dovedește însă foarte riscant prin anumite trăsături care provin chiar din caracteristicile Big Data:

– *Deși datele sunt anonime, totuși, la intersecția unor fluxuri detaliate de informații anonime devine posibilă dezanonimizarea acestora*. Dat fiind volumul mare și diversitatea surselor care compun Big Data, nivelul de detaliu la care sunt de-scrise punctele din norul de date (*data points*) face posibilă identificarea totală sau la un anumit nivel de generalitate/precizie a persoanelor implicate. Prin urmare, în studiile bazate pe Big Data, premisa trebuie să fie aceea că dezanonimizarea și inferarea unor informații noi și posibil dăunătoare despre subiecți umani individuali, fără nici o agregare statistică, este un risc real pentru toți cei implicați.

– Deși subiecții umani și-au dat acordul, direct sau implicit, pentru generarea acestor date (fie conform legii, fie prin Termenii și condițiile unor servicii la care au optat etc.), aceștia adesea nu au anticipat că datele despre ei vor fi refolosite într-un context cu totul nou, cel al cercetării. *Decontextualizarea și recontextualizarea datelor este specifică Big Data și utilizărilor sale*.[57] Această trăsătură este semnificativă pentru etica științei și pentru etica utilizării datelor în general. Astfel, un tip de informații care au fost legitim culese și prelucrate într-un anumit context pot ajunge să fie transferate, analizate și folosite într-un alt context, în care legitimitatea inițială nu se mai aplică.

– Studiile realizate pe Big Data sunt adesea strâns legate de intervenții asupra populațiilor și chiar asupra indivizilor descriși de o astfel de resursă. În exemplul de mai sus, identificarea omului din spatele imaginii artistice Banksy ar avea consecințe clare și direct atribuibile, cel puțin parțial, cercetătorilor.

– Distanța dintre cercetători și subiecți, atât în termeni de spațiu, cât și de timp sau de context social, nu este o garanție a lipsei de impact negativ. Dimpotrivă, această distanță, precum și scuza caracterului public al datelor, îi poate face pe cercetători să fie neglijenți. Astfel, ei pot ajunge să ignore prejudiciile reale pe care le pot aduce indivizilor, dar și populațiilor descrise de studiile lor.

Un alt exemplu de studiu care ilustrează aceste riscuri cu implicații etice

se referă la analiza unei comunități de pe platforma de socializare Twitter, formată din simpatizanți ai mișcării teroriste Statul Islamic din Irak și Levant (ISIS).[58] Autorii studiului propun o metodă nouă de analiză a conversațiilor și rețelelor din platformă. Metoda permite, printre altele, detectarea celor mai influente voci care promovează mesajele organizației teroriste. Articolul analizează un lot de 119.156 de conturi incluse în studiu, din care detectează o comunitate activă de susținători, precum și pe cei mai vocali membri în ceea ce privește activitatea și persuasiunea în rețea.

Acest studiu evidențiază trăsăturile menționate anterior, aferente cercetărilor bazate pe Big Data. Astfel, se pornește de la datele disponibile public (și postate într-un anumit context dat) ale utilizatorilor unei rețele de socializare, iar cercetătorii dezvoltă o metodă și propun o clasificare a acelor subiecți. Aceste acțiuni dezvoltate de echipa de cercetători au un potențial ridicat de a conduce la efecte directe negative asupra unor membri ai respectivei comunități, dat fiind asocierea cu o mișcare teroristă. Metoda propusă pornește de la date, le recontextualizează și înlesnește o intervenție asupra subiecților – fără să fi fost implicat vreun consimțământ al acestora.

Desigur, prevenirea sau sancționarea terorismului este o activitate virtuoasă, de unde derivă un tip de legitimare a acestui studiu. Totuși, remarcăm că noua metodă propusă de cercetători (IVCC – *Iterative Vertex Clustering and Classification*) poate fi transferată în contexte diverse, cu scopuri multiple – inclusiv pentru a preveni sau sancționa alte mișcări sociale. Ce legitimitate ar avea o astfel de cercetare aplicată asupra mișcării #rezist din România anilor 2017–2019, sau mișcării #BlackLivesMatter inițiată în SUA în 2013, în loc de Statul Islamic?[59]

Modelele tradiționale ale eticii cercetării au fost centrate pe păstrarea autonomiei decizionale individuale prin consimțământ informat, dar noile metodologii bazate pe studii ale rețelelor și pe date vaste redefinesc radical rolul individului și capacitatea sa de control asupra informațiilor pe care le oferă.[60] Este vorba aici atât despre subiectul datelor, respectiv subiectul cercetării, dar și despre cercetători. Aceștia din urmă au responsabilitatea de a proteja participanții la cercetare de posibile daune derivate pe parcursul sau prin utilizarea cercetării, dar devine din ce în ce mai greu de anticipat cum vor fi folosite rezultatele produse de ei.

La un cu totul alt nivel de risc, cercetarea în domeniul recunoașterii faciale și a detectării diferitelor trăsături personale pe baza imaginii feței poate fi asociată cu încălcări grave ale drepturilor omului. De exemplu, un studiu finanțat de guvernul chinez, publicat în 2018, prezintă o soluție de identificare a etniei pe baza informațiilor obținute prin detectarea facială automată, cu exemplificare asupra etniei uigure.[61] Această publicație apare în contextul în care guvernul chinez utilizează soluții de recunoaștere facială și de detectare a etnicității pentru supravegherea continuă a minorității uigure din provincia Xinjiang. Cetățenii acestei etnii suferă de grave încălcări ale drepturilor omului, atât la nivelul încălcării libertăților umane fundamentale, cât și datorită forțării lor de a participa la procese de reeducare ce au loc în lagăre de detenție.[62] Comunitatea academică a protestat, solicitând clarificări cu privire la obținerea consimțământului participanților. Chiar dacă participanții și-au oferit în mod liber și asumat consimțământul (ceea ce nu este ușor de confirmat într-un regim nedemocratic), riscurile de agravare a încălcării unor libertăți fundamentale prin aplicarea unei astfel de soluții bazate pe Big Data și învățare automată sunt ușor de imaginat.

În concluzie, studiile bazate pe Big Data ridică noi provocări pentru etica științei. Chiar acele trăsături ce par a garanta siguranța participanților și detașarea lor de vreun posibil impact negativ al studiilor, precum anonimitatea datelor, volumul vast și prelucrarea statistică sau distanța spațială și temporală față de cercetători, sunt, de fapt, surse de vulnerabilitate și de risc. În condițiile în care Big Data sunt persistente și traversează contexte diferite de colectare și utilizare, *este necesară regândirea formelor de evaluare etică a cercetărilor și de autoreglementare a deontologiei în domeniul profesional al cercetării științifice, nu numai în zona academică, dar și în mediul privat și în cel public*. Dacă rămânem în registrul tradițional al reglementărilor bazate pe riscurile studiilor experimentale din zona clinică, vom ignora potențialele vătămări și afectări ale autonomiei oamenilor descriși de studiile bazate pe date vaste. Mulți dintre noi generăm cantități semnificative de informație în mediul online, iar datele noastre au fost și vor mai fi, probabil, incluse în cercetări științifice. Acordul implicit, în astfel de cazuri, este ca rezultatele studiilor

să nu contravină drepturilor noastre fundamentale, intereselor noastre legitime sau interesului public. Evaluările etice ale ştiinţei datelor trebuie să aibă în vedere conexiunea dintre numere şi persoane, pentru a evita delegitimarea morală a demersurilor ştiinţifice.

Deficitul de adevăr factual: Big Data, ştirile false şi infodemia

Gura omului sparge cetăţi.

Secţiunea dedicată problematizării etice a deficitului de cunoaştere din soluţiile Big Data nu se poate finaliza fără o discuţie despre manifestarea sa cea mai vizibilă, şi anume ştirile false (fake news). Proliferarea informaţiilor false sau parţiale, fie cu intenţie malefică, fie animată de alte intenţii, din surse dintre cele mai diverse, a devenit o trăsătură distinctivă a sferei digitale a zilelor noastre. Ştirile false poluează mai ales circuitele din zona reţelelor de socializare, punând la îndoială capacitatea internetului de a funcţiona ca o amplificare a cogniţiei umane[63]. În contextul pandemiei Covid-19, vorbim despre o infodemie, în care transmiterea virală a patogenilor este însoţită şi agravată de transmiterea virală a falselor informaţii privind cauzele şi soluţiile acestei dezordini.[64]

În 2016, jurnaliştii de la *BuzzFeed News* au făcut o descoperire surprinzătoare. Una dintre localităţile cele mai fervent implicate în propaganda electorală americană, cu orientare favorabilă candidatului Donald Trump, era oraşul Veles, de aproximativ 45.000 de locuitori, din Republica Macedoniei de Nord.[65] Care să fie sursa acestui partizanat neaşteptat? Investigaţia la faţa locului a lămurit situaţia. Adolescenţi şi tineri din Veles creaseră peste o sută de site-uri destinate politicii americane, cu titluri precum WorldPoliticus.com, TrumpVision365.com, USConservativeToday.com, care publicau, aproape integral, ştiri false cu conţinut pro-republican. Tinerii nu erau ei înşişi simpatizanţi ai vreunui partid politic sau candidat, dar apreciau sincer banii câştigaţi din reclame. Accesările pe care le cumulau (prin Google sau Facebook) pentru noutăţile publicate generau venit pentru platforme prin intermediul plasării de reclame. Fracţiunea care ajungea la ei a fost suficientă pentru a motiva crearea unei întregi industrii propagandistice. Câteva exemple de titluri

propagate de adolescenții macedoneni pe frontul politic american includeau: „Papa Francisc le interzice catolicilor să voteze pentru Hillary!“; „Oprah declară la FOX News că «Unii oameni albi trebuie să moară»; raționamentul său este incredibil!“; „Ultimă oră: există dovezi că Obama s-a născut în Kenya; Trump a avut dreptate de la bun început...“ Tinerii experimentaseră și cu conținut pro-Bernie Sanders, dar, pur și simplu, nu fusese la fel de profitabil. Propaganda falsă pro- Trump se dovedise o mină de aur. Vedem, în acest caz, cum știrile false apar prin complicitatea creatorilor de conținut și a platformelor, într-un joc în care ambii câștigă, iar publicul pierde. Atrăgând atenția cititorilor dornici de emoții tari și de confirmarea dorințelor politice, platformele țintesc reclame, iar emițătorii de mesaje atractive sunt recompensați pentru momelile plasate. Adevărul comunicării sau consecințele sale politice sunt complet irelevante pentru acest spectacol al capitalismului datelor.

Conceptul de *fake news* (știri false) a avut o ascensiune foarte interesantă în conștiința colectivă. În figura 33, putem vedea popularitatea sa relativă în căutările pe Google din Statele Unite ale Americii, în perioada 2004–2021. Figura 34 arată evoluția similară la nivel global. Dacă termeni înrudiți, precum manipulare (*manipulation*) sau teoria conspirației (*conspiracy theory*) au o oarecare stabilitate în vocabularul public, sintagma *fake news* a intrat relativ recent sub lumina reflectoarelor. Ascensiunea sa este legată de niște *tweet*-uri ale președintelui Donald Trump. Prima creștere în vizibilitate a apărut în februarie 2017, când președintele american a acuzat mai multe trusturi media care îl portretizează negativ, că publică, de fapt, *fake news*[66]. În ianuarie 2018, a fost înregistrat al doilea vârf de popularitate al noțiunii, când președintele Trump a apelat la Twitter pentru a aplica eticheta de Fake News Awards pentru diferite știri și canale media.[67] Conceptul a rămas, astfel, ancorat în discursul public. În SUA însă acesta nu a mai devenit niciodată la fel de popular în căutări, având niște vârfuri ulterioare, mult mai mici, generate de pandemia Covid-19 (martie 2020) și de alegerile prezidențiale din noiembrie 2020. La nivel global însă, conceptul a fost relansat în circulație, cu și mai multă forță, în octombrie 2018, cu ocazia alegerilor prezidențiale din Brazilia[68], și în martie 2020, când infodemia era recunoscută ca fenomenul geamăn al pandemiei.[69]

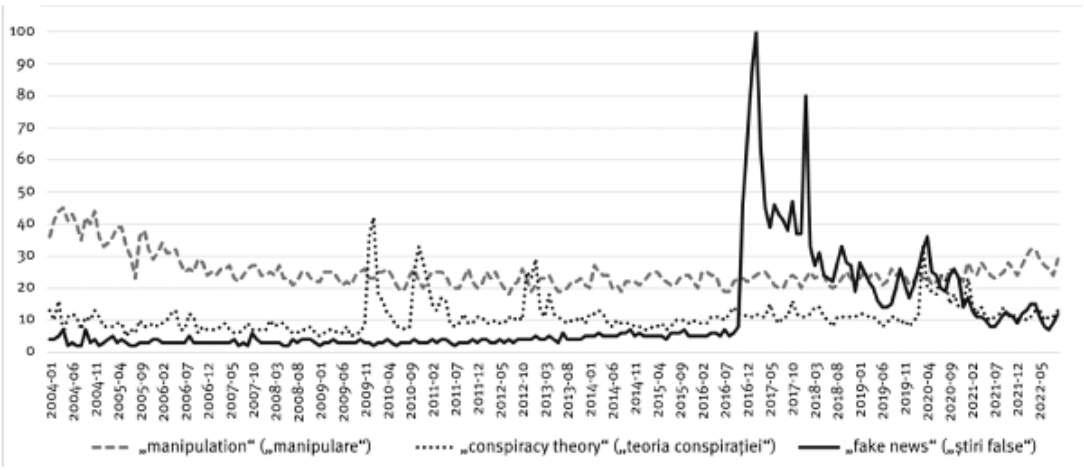


Fig. 33 Popularitatea relativă a termenilor „manipulation“, „fake news“ și „conspiracy theory“ în căutări cu motorul Google, în SUA, 2004–2021. Sursa: Google Trends.

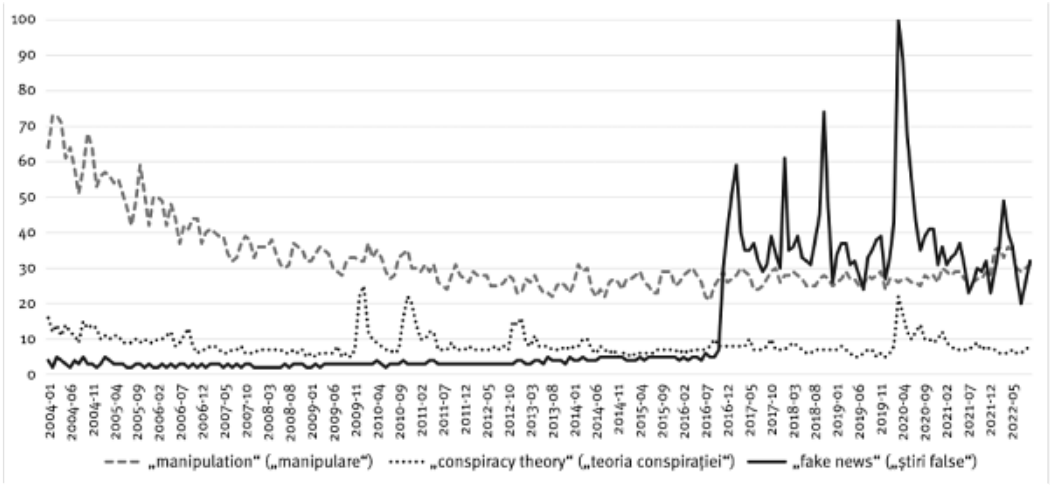


Fig. 34 Popularitatea relativă a termenilor „manipulation“, „fake news“ și „conspiracy theory“ în căutări cu motorul Google, la nivel global, 2004–2021. Sursa: Google Trends.

Remarcăm că eticheta de *fake news* și-a câștigat popularitatea bidirecțional, ca acuzație reciprocă în controversele politice și epidemiologice ale ultimilor ani. Aceeași expresie este folosită cu denotații contrastante în discursuri diferite, ceea ce creează senzația că părțile opuse ale spectrului politic nici nu mai vorbesc aceeași limbă.[70]

Disponibilitatea tot mai crescută a unor versiuni personalizate ale

adevărului, care sunt pe placul oricărui consumator media, este legată și de *polarizarea politică*. Studiile realizate de Pew Research Center arată că, din 2004 până în 2017, în Statele Unite, s-a petrecut o intensificare continuă a partizanatului politic atât la nivelul publicului[71], cât și în mass-media consumată de alegători[72]. SUA are însă, și în această privință, un profil distinctiv în raport cu alte state democratice. Cel puțin în privința evaluării politicilor privind Covid-19, SUA prezintă un nivel semnificativ mai ridicat de polarizare decât țările europene, precum Germania, Italia sau Suedia – deși Franța, Spania și Regatul Unit au niveluri doar puțin mai scăzute comparativ cu Statele Unite.[73]

Dată fiind puterea pe care SUA o posedă în lansarea platformelor digitale folosite la nivel global, apare întrebarea dacă nu cumva aceste tehnologii exportă nu doar beneficiile digitalizării, ci și polarizarea politică din țara de origine. După cum am văzut prin controversele repetate în care a fost implicat Facebook – de la Cambridge Analytica la dezvăluirile lui Frances Haugen[74] –, această ipoteză pare tot mai plauzibilă. Chiar dacă tehnologiile promit neutralitate morală sau chiar o valență pozitivă, forma practică pe care o iau prin modelul lor de afaceri, adaptat tipului de societate din SUA, le transformă în instrumente foarte eficace în slujba actorilor politici, sociali sau comerciali interesați în captarea audienței cu orice preț, inclusiv în instrumente ale distorsionării realității și polarizării societății.

Agnosticismul algoritmilor de recomandare

Spre deosebire de deficitele de precizie, cunoaștere și înțelepciune, care sunt specifice modelelor ML/AI bazate pe date vaste, problematica Big Data este legată indirect de știrile false. Legătura apare prin intermediul consolidării internetului și rețelelor socializare, care sunt animate de industria marketingului personalizat în care Big Data și predicțiile pe baza profilurilor digitale joacă un rol esențial. Algoritmii de recomandare sunt antrenați pentru a ne stimula, a ne atrage și a ne păstra pe platforme, nu pentru a discerne adevărul de fals. Motto-ul modelelor de afaceri ale marilor platforme digitale pare să fie: „Nu lăsa faptele să stea în calea unei povești bune“.

Personalizarea fluxurilor informaționale și stimularea indivizilor pentru a deveni creatori și comunicatori de opinii, difuzate apoi în bule de conținut[75], sunt factori care au amplificat dimensiunile fenomenului infodemiei. *Legătura apare, prin urmare, mai ales datorită modelelor de afaceri ale platformelor bazate pe marketing personalizat*, care motivează aceste fragmentări ale orizonturilor de cunoaștere sociale în sub-lumi conectate digital.

După cum am văzut și în perioada pandemiei, știrile false pot conduce direct la daune umane și sociale. De la cazuri de oameni ce refuză tratamentul medical pentru infecția cu coronavirus din cauza unor panici alimentate digital și reticența față de vaccinare la scară largă la proliferarea argumentelor pseudoștiințifice împotriva purtării măștii, știrile false au condus la agravarea impactului pandemiei în termeni de vieți pierdute și persistență a dezordinii sociale și economice. Filtrele algoritmice de care depindem astăzi pentru a ne selecta informațiile din preaplinul internetului s-au dovedit ușor manipulabile de actorii sociali interesați să-și propage mesajele. Platformele, motivate comercial să caute noi utilizatori și surse de atracție, indiferent de valoarea lor de adevăr, au înlesnit transmiterea mesajelor captivante și emoționante, dincolo de vreo verificare a veridicității acestora. *Rețelele de socializare* din mediul online și modelele lor de business sunt responsabile, în mare măsură, de transmiterea mesajelor antivaccin, antimască și, în general, de proliferarea teoriilor conspiraționiste.[76] În același timp, după cum vom vedea mai jos, *motoarele de căutare* sunt și ele manipulabile, mai ales în momente de criză, prin injectarea de conținuturi false atunci când există un dezechilibru între cererea și oferta de informații.

Golurile de date și prăbușirea în lumi contrafactuale

Uneori, când călătorim cu avionul, ajungem într-un gol de aer și simțim cum ne prăbușim. Cele mai puternice avioane sunt vulnerabile la astfel de lacune în mediul în care funcționează. Pentru motoarele de căutare, una dintre sursele noastre esențiale de informație, acest mediu este compus din datele tuturor site-urilor publicate pe internet. Ce se întâmplă însă când,

legat de un anumit subiect, nu există suficiente site-uri?

Michael Golebiewski și Danah Boyd[77] au introdus conceptul de „goluri de date“ (*data voids*) pentru a explica felul în care rezultatele motoarelor de căutare sunt manipulabile prin injectarea de informații false formulate atent și la momentul potrivit. Motoarele de căutare funcționează prin ierarhizarea miilor sau milioanelor de site-uri care oferă răspunsuri relevante la căutările noastre. Problema apare atunci când utilizatorii folosesc combinații rare sau demodate de cuvinte-cheie, pentru care nu există prea multe informații online. De exemplu, dacă vom căuta, în limba engleză, pe Google Search, *robotic enemies* („dușmani robotici“), obținem 61 de milioane de rezultate în 0.61 secunde. Dacă trecem de la căutarea în engleză la căutarea în română, pâlnia se îngustează vizibil. Pentru cuvintele-cheie „dușmani robotici“ ajungem la doar 47.000 de rezultate. Dacă, apoi, îngustăm la „dușmani robotici ai neamului pisicesc“, Google Search intră în impas total și nu găsește nimic relevant. Acesta este un exercițiu pur teoretic, dar situații similare pot apărea în realitate, cu mize mult mai importante.

Golebiewski și Boyd relatează un eveniment tragic petrecut în data de 5 noiembrie 2017 în Sutherland Springs, Texas. Astăzi, dacă vom căuta pe Google Search acest nume, obținem peste patru milioane de rezultate, dintre care primul este pagina de Wikipedia pe tema „Sutherland Springs Church Shooting“ („Atacul armat de la biserica din Sutherland Springs“).[78] Înainte de această zi însă, numele orășelului era rareori căutat și nu existau prea multe site-uri care să facă trimitere la aceste cuvinte-cheie, dincolo de indicații rutiere și predicții meteorologice. În ziua de 4 noiembrie 2017, un bărbat înarmat cu o armă semiautomată a intrat în biserica baptistă din Sutherland Springs și a omorât 26 de oameni, rănind alți 22, înainte de a se sinucide. Dintr-odată, publicul american a început să caute fervent știri pe această temă. În acest gol de date, o rețea de dezinformare a împins pe rețelele sociale Twitter și Reddit știrea falsă că atacatorul ar fi fost membru al Antifa, o rețea descentralizată de activiști antifasciști din SUA.[79] Acest efort propagandistic urmărea influențarea opiniei publice prin învinovățirea unei mișcări stângiste pentru această tragedie. Dacă Antifa putea fi poziționată ca țap ispășitor, atenția ar fi fost distrasă de la rolul regimului armelor în epidemia de omoruri în masă din SUA.[80] Pe măsură ce lucrurile

s-au clarificat, în mod paradoxal, problema a fost amplificată de chiar investigația care a demascat această campanie de dezinformare. Un articol publicat în *Newsweek* diagnostica situația, purtând titlul: „«Antifa» Responsible for Sutherland Springs Murders, According to Far-Right Media“ („«Antifa » este responsabilă pentru omorurile din Sutherland Springs, conform presei de extremă dreaptă“).[81] Cum rezultatele motoarelor de căutare preiau doar primele cuvinte din rezultate, în primele ore ale evenimentului cei interesați vedeau, la loc de frunte, informația trunchiată: „«Antifa» Responsible for Sutherland Springs Murders...“

Acest eveniment arată puterea pe care o rețea de emițători o poate avea în manipularea opiniei publice, intervenind imediat după ce o știre de ultimă oră propulsează în atenția cititorilor și spectatorilor un șir de cuvinte-cheie pentru care există un gol de date. Mecanisme similare pot fi folosite și pentru a contribui la curente de opinie mai stabile. În aceeași analiză, Golebiewski și Boyd arată cum a fost poziționată în discursul public ideea că atacurile armate din școlile americane sunt, de fapt, evenimente înscenate de la cap la coadă, prin care politicienii de stânga doresc să atace legalitatea portului de armă. Inovația a constat în generarea unor site-uri dedicate conceptului de *crisis actors* („actori de criză“), o sintagmă rareori căutată anterior. O întreagă construcție informațională a portretizat victimele acestor atacuri, inclusiv supraviețuitorii și părinții, ca fiind, de fapt, actori plătiți pentru a juca rolul unor persoane împușcate sau traumatizate. În 2018, după masacrul de la liceul Marjory Stoneman Douglas[82], un reporter a încercat să demaște această dezinformare, întrebându-l pe un elev supraviețuitor al atacului dacă este, cumva, un actor de criză. Publicul a fost astfel expus în masă acestui concept nou. Cei care au încercat să se lămurească, folosind căutarea pe internet, au ajuns direct în capcană. În lipsa altor date pe această temă, căutările pe Google după sintagma „crisis actors“ conduceau direct la site-urile conspirații oniste asociate expresiei, promovând astfel propaganda ce nega realitatea acestor tragedii.

Tactici similare sunt responsabile de popularizarea unor știri false prin motoarele de căutare și algoritmii de recomandare din rețelele de socializare, atunci când anumite expresii devin brusc populare. În lipsa unor

asocieri între aceste expresii și informații corecte disponibile online, apare o oportunitate de influențare a opiniei publice. Manipulatorii pot interveni rapid și concertat, direcționând motoarele de căutare spre site-uri optimizate special pentru a oferi răspunsuri care servesc unei anumite agende politice sau comerciale, fără legătură cu adevărul lor factual.

Discuțiile privind știrile false conectează un preaplin de relevanță cu un deficit de veridicitate. Tehnologiile informației au fost transformate în produse foarte profitabile, oferind oamenilor ceea ce vor să găsească – fie că este vorba despre pantofi, vremea probabilă sau povești despre candidații politici ai momentului. În ceea ce privește moda, gusturile nu se discută și, oricât de impermeabile ar fi filtrele noastre informaționale, nu riscăm decât să trecem de la sublim la ridicol. În ceea ce privește descrierea societății însă, deliberarea este cheia găsirii răspunsurilor celor mai apropiate de adevăr. Preferințele noastre nu mai sunt suverane, deoarece faptele nu sunt opționale. Accentul pe care noile infrastructuri ale cunoașterii îl pun pe vești captivante, pe fluxuri rapide de informații și pe reacții emoționale puternice, transmise la scară largă, a transformat factualitatea într-o chestiune de gust. Atunci când căutăm răspunsuri relevante pentru noi, personal, platformele energizate de marketing predictiv riscă să ne ofere doar răspunsurile pe care vrem să le auzim. Producătorii și consumatorii de răspunsuri sunt prinși într-un joc de-a șoarecele și pisica, fugărindu-se prin labirintul algoritmic al adevărului, în urmărirea propriilor interese. Ceea ce putem observa cu claritate este că abundența de informație nu este sinonimă cu buna informare. Aceleași tehnologii care ne răspund în microsecunde la orice curiozitate sunt și cele care ne pot conduce spre ficțiuni construite în pofida realității, care ne sunt oferite pentru a ne captiva și a ne capta, în scopuri politice sau comerciale. Această vulnerabilitate a fluxurilor informaționale nu este însă nici uniformă, nici constantă și nici inevitabilă. În timp, platformele și motoarele de căutare își adaptează strategiile de filtrare și recomandare, căutând să-și păstreze reputația de surse credibile de informație și reacționând la sancțiunile impuse de societățile-gazdă. De aceea, este important să ne întrebăm care este echilibrul pe care îl dorim, la nivel social, între logica urmăririi preferințelor și logica urmăririi adevărului.

Deficienţe adversariale: războiul cunoaşterii în societatea algoritmică

Rău cu rău, dar mai rău fără rău...?

O sursă sistematică de erori şi imprecizii în capitalismul datelor poate apărea chiar din natura competitivă a acestuia. Oponenţii politici sau comerciali folosesc tot mai mult strategii de predicţie a comportamentului indivizilor atât pentru câştigarea atenţiei votanţilor sau consumatorilor, cât şi pentru subminarea adversarilor. Atacurile cibernetice, motivate militar, financiar sau politic, introduc noi surse de dezordine. Într-o luptă a tuturor contra tuturor, adversarii creează zgomot informaţional la niveluri greu de estimat.

Un exemplu interesant constă în aşa-numita fraudă a clickurilor (*click fraud*). În momentul în care accesările sau clickurile se convertesc în bani, apar şi actori care folosesc tehnologia pentru a produce false impresii şi acţiuni, imitând utilizatori umani şi pretinzând recompensă financiară pentru acţiunile lor. Estimările pagubelor cauzate de fraudele din AdTech sunt de ordinul miliardelor de dolari anual: 6,5 miliarde în 2017[83], respectiv 35 de miliarde în 2020[84].

Strategiile de fraudă digitală în zona AdTech sunt extraordinar de diverse.[85] Programe maliţioase pot suprapune reclamele, astfel încât o singură accesare să se numere ca fiind mai multe (*ad stacking*). Reclamele pot fi comprimate în imagini minuscule, apărând pe ecranul nostru cu sutele, fără ca noi să observăm şi să sesizăm, dar generând accesări plătite (*pixel stuffing*). Aplicaţii maliţioase cu aparenţă inofensivă (un ceas deşteptător, un calendar) pot rula continuu reclame invizibile în spatele ecranului, pentru a obţine banii firmelor fără ca utilizatorii finali să fie conştienţi de mesajele pe care, teoretic, le văd (*click injection*). Unele aplicaţii maliţioase pot anticipa clickurile noastre, sărind în faţă şi furând banii de la aplicaţia legitimă pe care o folosim (*install hijacking*). Există numeroase metode prin care boţii sau aşa-numitele ferme de clickuri, în spatele cărora se află oameni din Bangladesh, China, India, Indonezia şi alte ţări cu forţă de muncă ieftină, pot pretinde că sunt utilizatori umani legitimi. Aceştia pot vizualiza sau face click pe sute sau mii de reclame, în timpul

tipic alocat unui singur click (*click spamming*). Totodată, boții și astfel de imitatori profesioniști devin tot mai eficienți în a evita detecția.[86]

Pe măsură ce devin mai sofisticați, boții malițioși pot ajunge chiar să completeze chestionare online sau aplicații complicate cu date ale unor indivizi reali, extrase din profiluri digitale. Imaginați-vă că o universitate încearcă să recruteze studenți. Aceasta apelează la mai multe firme, pentru a atrage absolvenții de liceu să depună dosare de admitere la programele sale. Universitatea va plăti firmele pentru fiecare dosar depus prin intermediul lor. Inevitabil, nu toți studenții care depun dosarul și sunt admiși vor confirma înscrierea finală, deoarece mulți candidează simultan la mai multe programe. Universitatea își asumă riscul de a atrage candidați care, în cele din urmă, vor decide să meargă în altă parte. Totuși, ce se întâmplă dacă unii dintre așa-numiții candidați nu sunt, de fapt, reali? Costurile pentru universitate vor fi mult mai mari decât inițial anticipate. Unele firme pot genera dosare fictive, cu ajutorul boților care compun aplicații credibile, combinând date ale unor indivizi reali. Un sfert de licean din Texas, un sfert de funcționar din New York și jumătate de farmacist din California pot oferi datele personale necesare completării unui dosar de candidatură. Acest lucru se întâmplă deja în Statele Unite. Recrutorii universităților din SUA resping din start peste 50% din dosare, deoarece sunt false, adică sunt create cu ajutorul boților pentru a câștiga banii oferiți pentru dosare reale depuse.[87]

Desigur, frauda este combătută tot cu ajutorul boților, care verifică mesajele și linkurile pentru a evita atacurile malițioase, curăță bazele de date și realizează nenumărate operațiuni de monitorizare, verificare, arhivare și menținere a coerenței informațiilor. Avem, așadar, o competiție între boți atacatori și boți detectori, între surse de zgomot și semnale, în care puterea automatizării și inteligenței artificiale creează simultan cantități vaste de informație și eroare.

Complexitatea ridicată a modelelor de inteligență artificială le face să fie vulnerabile la manipulări subtile. S-a dezvoltat chiar un domeniu de studiu dedicat acestor atacuri, denumit „învățare automată adversarială“ (*adversarial machine learning*). Atacurile pot interveni în oricare dintre etapele modelării și deciziei. Un sistem de clasificare algoritmică este antrenat pe un set de date de învățare, din care deprinde niște reguli pe care

le aplică apoi pe datele reale, cele care trebuie analizate. Conor O'Sullivan discută trei tipuri de ofensive împotriva sistemelor de inteligență artificială, în funcție de momentul intruziunii.[88] *Otrăvirea datelor* se petrece atunci când atacatorii poluează datele de antrenare ale modelului, încercând să-i inducă anumite criterii decizionale. De exemplu, un sistem poate ajunge la concluzia (greșită) că studenții cu o combinație specifică de note la bacalaureat (să zicem, 8,80 la română, 7,95 la matematică, 7 la latină și 6 la morală) sunt cei mai buni candidați pentru admiterea la universitate. Această înșelătorie poate fi pusă în practică dacă atacatorii reușesc să dreseze modelul, strecurând astfel de cazuri fictive în datele de antrenare și marcându-le cu eticheta „admis". Intrigile caragialești tip „Lanțul slăbiciunilor" se mută astfel din registrul rugăminților umane în cel al persuasiunii algoritmice. „Să nu zici că nu poți!... Știu că poți! trebuie să poți!" insistă povestitorul în „Bacalaureat"[89]. Pentru a convinge un model decizional algoritmic să ne facă pe plac, sunt necesare alte abordări. Un atacator ar putea să păcălească și un model gata antrenat, coafând niște date ilegitime pentru a părea legitime. Acest tip de intervenție poartă denumirea de *evaziune la nivelul datelor*. În exemplul de mai sus, sistemul ar primi un dosar falsificat al candidatului, cu note mai mari la bacalaureat decât a avut în realitate. Să ne gândim la o altă situație, în care cineva dorește să trimită mesaje comerciale în masă, mesaje ce ar fi, probabil, etichetate ca fiind spam. Cum pot fi strecurate astfel de oferte printre cerberii filtrelor algoritmice? Atacatorii vor experimenta continuu, tot cu ajutorul modelelor automate, până vor găsi anumite combinații foarte specifice de cuvinte care păcălesc sistemul. Astfel, prin încercare și eroare, vor găsi cuvinte pe care să le înlocuiască cu sinonime, pentru ca un mesaj spam să nu mai fie clasificat ca atare. Acestea sunt așa-numitele „atacuri prin parafrazare"[90], care nu vor fi detectate de un cititor uman, deoarece sensul rămâne același. Este posibil să fie suficient ca, în loc de „Vă rugăm să ne contactați pentru oferte de credit incredibil de avantajoase", să scriem „Vă rugăm să ne apelați pentru oferte de credit incredibil de atractive". Dacă e vorba despre procesarea imaginilor, atacatorii pot să modifice câțiva pixeli din poza unui panda, pentru ca acesta să pară un gibon, fără ca transformarea să fie vizibilă pentru un om, observă O'Sullivan. Prin urmare, nu e de mirare că

mașinile autonome pot ajunge să creadă că fulgii de zăpadă, umbrele copacilor sau pescărușii sunt obstacole de neocolit.[91] Al treilea tip de atac se referă la *furarea modelului* cu totul, pentru a infera criteriile folosite sau pentru a extrage din el date prețioase. Dacă au modelul la îndemână, atacatorii îl pot testa îndelung, pentru a afla cum funcționează sau pe ce informații a fost antrenat. Progresele rapide în învățarea automată adversarială conduc la o preocupare tot mai ridicată pentru imunizarea modelelor ML/AI în fața unor astfel de intruziuni. Dincolo de ecranele care ne livrează mesaje și decizii, se petrece un joc de poker continuu, imperceptibil cu ochiul liber, în care algoritmii încearcă neîncetat să se înșele și să-și citească adversarii după cele mai mici indicii.

Lupta împotriva știrilor false, adesea propagate cu ajutorul boților[92], a ajuns de asemenea să fie din ce în ce mai mult purtată cu ajutorul inteligenței artificiale. În secțiunea anterioară, am văzut cum rețelele hiperconectate, în care știrile curg pe fluxul emoțiilor colective, pot conduce la popularizarea unor versiuni contrafactuale ale lumii. Cum pot fi găsiți arbitri ai rezonabilității, în această lume în care viralitatea mesajelor este monetizată atât de emițători, cât și de platformele care fac posibilă comunicarea? Facebook promite o soluție tehnologică la problema dezinformării. În încercarea de a-și păstra modelul de afaceri, ce s-a dovedit atât de profitabil, gigantul tehnologic construiește versiuni tot mai ample de filtre algoritmice care să poată detecta și sancționa automat știrile false. În decembrie 2021, după dezvăluirile lui Frances Haugen privind eșecurile companiei în prevenirea discursurilor incitatoare la ură și a știrilor false (vezi capitolul 1), Facebook a anunțat implementarea unui model de detecție dintr-o nouă generație.[93] Acest sistem este numit Few-Shot Learner, adică „cel ce învață din puține încercări“. Spre deosebire de modelele anterioare, care aveau nevoie de foarte multe exemple pentru a învăța ce este corect și ce este incorect, acesta combină un antrenament generic pe cantități enorme de postări cu un antrenament specific, mult mai țintit și mai rapid. Promisiunea companiei este că va putea preveni difuzarea știrilor false în peste o sută de limbi, reacționând în câteva săptămâni, și nu în câteva luni. Facebook nu este singurul care inovează: emițătorii mesajelor se vor adapta, la rândul lor, la noile filtre și proceduri, automatizându-și

campaniile de comunicare cu noi instrumente și tactici. Așadar, rămâne de văzut cum va fi pusă în practică această promisiune și în ce măsură platforma va putea diferenția, cu discernământ și la scară globală, mesajele informative de cele manipulative.

Probabil că exemplul cel mai clar de conflict între sisteme adversariale de inteligență artificială este chiar războiul. În decembrie 2021, cele 125 de state participante la *Convenția Națiunilor Unite privind anumite categorii de arme convenționale* s-au reunit la Geneva. Pe agenda discuției s-a aflat și propunerea de interzicere a armelor autonome, numite și „roboți ucigași“ (*killer bots* sau *slaughterbots*). Acestea sunt sisteme care declanșează atacuri cu potențial letal asupra țintelor, fără intervenția unui supervizor uman, folosind sisteme de recunoaștere facială și alte procese decizionale bazate pe IA. Un raport al Consiliului de Securitate al Națiunilor Unite a semnalat prima utilizare a unor arme autonome în războiul din Libia în martie 2020, și anume dronele kamikaze STM Kargu, produse în Turcia.[94] Aceste drone sunt încărcate cu explozibil. Odată lansate, ele nu mai păstrează conexiunea cu atacatorii umani, ci își urmăresc ținta, o detectează și se aruncă asupra ei în mod autonom. Acest tip de atac este denumit „lansează-uită-găsește“ (*fire, forget and find*), deoarece, după ce este lansat, botul își uită originile și găsește singur ținta atacului.

Discuția privind excesul și deficitul cunoașterii bazate pe inteligență artificială și date vaste capătă cu totul alte mize atunci când vorbim despre sisteme automatizate menite să ucidă. Am văzut cum o eroare de identificare a unei fețe umane poate conduce la arestări abuzive, la imposibilitatea autentificării într-o aplicație sau la refuzul unor servicii sociale.[95] În acest caz însă, eroarea se poate solda direct cu pierderea vieților omenești, ajungând până la crime de război. James Dawes, specialist în drepturile omului și profesor la Macalester College, identifică mai multe surse de risc în ceea ce privește armele autonome.[96] *Erorile computaționale* ale modelului sunt primul nivel de vulnerabilitate, amplificat de lipsa de transparență ale acestor produse. Tehnologiile de inteligență artificială sunt și mai opace în mediul militar decât în cel comercial sau politic. Vor putea dronele să facă diferența între copiii care se joacă de-a războiul și inamicii înarmați cu arme reale? Vor putea să distingă civilii care fug din calea

luptelor de trupele implicate într-o retragere tactică? *Scara* la care operează aceste sisteme poate face ca anumite erori să se propage mult mai rapid și mai amplu decât ar fi făcut acțiunea umană. Dacă boții își pierd conexiunea cu comandamentul-sursă, erorile, odată detectate, nici nu mai pot fi corectate. Dincolo de riscurile erorilor de identificare și de țintire, Dawes discută și riscurile *proliferării*, distingând două versiuni. *Proliferarea inferioară* se referă la dezvoltarea piețelor de roboți ucigași tot mai ieftini, ce ar putea fi convenabil folosiți de grupările criminale organizate împotriva adversarilor – inclusiv jurnaliști, judecători sau alți oponenți.[97] Proliferarea superioară se referă la dezvoltarea de către actorii politici, statali sau comerciali cu resurse vaste a unor arme autonome foarte sofisticate și scumpe. Aceste arme vor redefini complet costurile și beneficiile războiului. În lipsa unui acord global de restrângere severă a uciderii automatizate, puterea de a ținti și puterea de a greși a sistemelor ML/AI va ajunge într-o cu totul altă arie de impact. SUA, Rusia și India, care investesc masiv în aceste tehnologii, s-au opus unei decizii ferme a Convenției de la Geneva privind roboții ucigași.[98] Până la următoarea reuniune a statelor semnatare, în 2026, omenirea va avea la dispoziție cinci ani în care să experimenteze, pe propria piele, implicațiile acestui nou fel de a transpune cunoașterea în acțiune.

---

12. Textul GDPR este disponibil la adresa: https://gdpr-info.eu.

13. Articolul 22 GDPR este disponibil la: https://gdpr-info.eu/art-22-gdpr.

5. În balanță: noi riscuri morale ale științei și societății algoritmice

Nu în mult stă bunul, ci în bun stă multul.

Societatea bazată pe decizii algoritmice și date vaste conduce la reconfigurări radicale ale cunoașterii și ignoranței în raport cu distribuțiile epocilor anterioare. Este necesară o problematizare etică a utilizării inteligenței artificiale și Big Data, mai ales a fluxurilor de date personale tranzacționate și agregate, raportată la reconfigurarea deficitului și excesului de cunoaștere.

Examinând dezbaterile publice curente, ne putem întreba care sunt ariile de eroare și incompletitudine informațională și care sunt zonele de exces sau asimetrie.

Deciziile algoritmice bazate pe Big Data sunt acuzate simultan de deficit și de exces cognitiv. Contradicția este doar parțială, fiind cea mai acută în ceea ce privește potențialul lor de a genera predicții precise ale preferințelor individuale și de a motiva, sau chiar de a manipula, indivizii să acționeze în interesul celor care controlează datele. Nu intră în sfera de discuție a acestei cărți să adjudecăm definitiv problema potențialului predictiv și manipulator al soluțiilor bazate pe Big Data. Este însă important să remarcăm relevanța acestei dezbateri în discuția publică, deoarece atinge direct valoarea libertății umane.

Un cititor anonim a trimis revistei *Wired* următoarea întrebare[1]: „Una dintre aplicațiile mele de streaming de muzică îmi creează liste personalizate de melodii și se pricepe înspăimântător de bine să prezică ce muzică o să-mi placă. Să înțeleg de-aici că sunt o persoană plictisitoare?“ În cultura individualistă occidentală, să fii predictibil înseamnă să fii neinteresant, plat, lipsit de creativitate și fără un nucleu autentic personal. Vrem să gândim neconvențional, lateral, critic, spontan și să apărem celorlalți, dar și nouă înșine, ca niște mistere profunde. Problema este că, în ziua de azi, tehnologia ne face pe toți să fim sau, cel puțin, să părem

predictibili. Noi înşine contribuim la această domesticire a misterului, oferind constant algoritmilor indicii şi feedback privind preferinţele noastre: de la clickuri şi like-uri la cumpărături, călătorii şi emoţii transmise prin expresii faciale sau vocale. Pe măsură ce modelele ML/AI ne vor prezice tot mai precis ce ascultăm, ce mâncăm, unde mergem, cu cine votăm, unde muncim şi dacă ne îmbolnăvim, ne vom spori oare controlul asupra destinului, sau ne vom încredinţa unei sorţi stabilite algoritmic?

Aşadar, când ne raportăm la excesul cognitiv, principala acuzaţie adusă produselor bazate pe Big Data şi ML/AI se referă la *degradarea autonomiei* atât la nivel individual, cât şi la nivel colectiv. *Autonomia individuală* este afectată de produsele predictive realizate prin reclame şi mesaje electorale, poziţionate în micromomente de vulnerabilitate decizională cu scopul maximizării şanselor de conversie. Anticipările algoritmice nu se limitează însă doar la reclame. În 2019, directorul executiv al IBM, Ginny Rometti, afirma că inteligenţa artificială poate prezice cu o precizie de 95% ce angajaţi îşi caută un alt loc de muncă.[2] Ce înseamnă, pentru libertatea noastră, faptul că organizaţia în care muncim ne cunoaşte planurile de viitor, chiar şi atunci când nu vrem să fie ştiute? Ce consecinţe ar avea un sistem continuu de detectare a emoţiilor angajaţilor pornind de la expresiile faciale, de la mesaje şi de la tonul vocii?[3] Mai mult, ce consecinţe ar avea asupra valorii vieţii noastre o predicţie tot mai precisă a momentului în care vom muri? Algoritmii ML/AI sunt angrenaţi în calcule tot mai complicate ale riscurilor de sănătate induse de stilul nostru de viaţă, de bolile cronice pe care le avem şi de soarta pe care, statistic vorbind, o au oamenii consideraţi similari nouă[4]. Dacă fiecare dintre noi va avea o etichetă digitală indicând, aproximativ, anii care ne-au mai rămas de trăit, cum vom fi reevaluaţi în calculele companiilor de asigurări, angajatorilor sau aplicaţiilor matrimoniale?

Discursul reducţionist al ştiinţei datelor, în care indivizii devin simple puncte în nori statistici sau noduri în reţea, foarte puţin relevante pentru imaginea de ansamblu, conduce la o degradare a conexiunilor dintre ştiinţa descriptivă a societăţii şi filozofia morală şi politică în care drepturile individuale au un rol central, şi individualitatea este o sursă de valoare, nu o sursă de eroare. Persistenţa informaţiilor în mediul digital poate fi, de

asemenea, o sursă de scădere a libertăţii individuale, afectând reputaţia indivizilor mult dincolo de contextele temporale considerate legitime la nivel social.

Autonomia colectivă este şi ea afectată prin vulnerabilizarea democraţiei, slăbită mai ales în momentele cruciale ale campaniilor electorale şi ale alegerilor, prin polarizare şi micro-ţintire electorală, dar şi prin infodemia de ştiri false. Utilizarea ilegitimă a datelor cu potenţial discriminatoriu afectează atât autonomia individuală, cât şi pe cea a categoriilor sociale afectate de tratament inegal; apar aici şi bucle de feedback prin care o definiţie iniţială a situaţiei ajunge să fie reconfirmată, dacă rezultatele unei selecţii discriminatorii se transformă în date pentru următoarele decizii algoritmice.

Nu în ultimul rând, *costurile financiare şi psihologice* ridicate ale proceselor bazate pe Big Data, distribuite inegal, acutizează asimetria de informaţii şi de putere. Datele vaste şi soluţiile ML/AI au, din cauza dimensiunilor lor tot mai ample, şi un cost ecologic care nu poate fi neglijat.

Dacă pornim de la *deficitul epistemic*, sursa problemelor etice apare atunci când o cunoaştere deficitară devine bază pentru acţiune. Big Data şi produsele bazate pe ML/AI sunt tot mai frecvent folosite, atât în organizaţii comerciale, cât şi în organizaţii publice, fiind implicate în procese de evaluare a populaţiilor la scară largă şi în deciziile rezultate pe baza acestor evaluări. Big Data sunt resurse ale sortării oamenilor, în decizii medicale, poliţieneşti, educaţionale, financiare, de angajare etc. Atunci când aceste acţiuni sunt bazate pe o cunoaştere defectuoasă, pot apărea probleme multiple, adresate polarizat în dezbaterea publică.

O primă problemă constă în *riscul luării unor decizii arbitrarii şi nejustificate*, care vor fi inechitabile şi discriminatorii pentru anumite persoane sau actori sociali. În ultimul deceniu s-a dezvoltat o problematizare amplă a părtinirii şi discriminării algoritmice, precum şi o înţelegere nuanţată a surselor sale (una dintre ele constând în calitatea slabă a datelor vaste care susţin deciziile ML/AI) şi a modalităţilor prin care poate fi identificată, monitorizată şi prevenită.

O a doua problemă are în vedere *posibilele daune aduse unor persoane* care se văd păgubite în raport cu aşteptările lor rezonabile, conform cărora

au luat decizii şi au făcut investiţii anterioare. Astfel de daune pot apărea fie datorită discriminării algoritmice, enunţate anterior, fie datorită unor erori aleatorii din date sau a unor buguri din software-ul ce procesează aceste date, sau datorită unor procese de modelare din cutia neagră a produselor ML/AI. Datorită complexităţii ridicate a modelelor bazate pe Big Data, este adesea foarte dificil de înţeles unde anume se află sursa unei decizii surprinzătoare sau dăunătoare.

O a treia problemă se referă la *degradarea încrederii consumatorilor şi a reputaţiilor companiilor*, afectând comunicarea şi feedbackul în relaţiile de afaceri. Nu doar indivizii pot fi afectaţi negativ de calitatea slabă a bazelor de date vaste, ci şi organizaţiile care apelează la ele pentru a segmenta publicul-ţintă şi pentru a-l microţinti cu reclame, anunţuri, politici sau oferte comerciale. Dacă această segmentare este eronată, organizaţiile se vor lovi de resentimentul publicului cu care interacţionează, pierzându-i încrederea.

O a patra problemă dezbătută public constă în *degradarea proceselor decizionale*, care pierd din înţelepciunea practică (*phronesis*) necesară adecvării la context şi la idealurile unei vieţi şi comunităţi virtuoase, devenind mecanice şi obscure. Big Data sunt esenţialmente descriptive, bazate pe modele care sintetizează şi simplifică datele prin identificarea de corelaţii. Predicţiile pe baza acestor modele riscă să devină politici conservatoare de reproducere a tiparelor sociale trecute, inclusiv a structurilor de putere şi de opresiune, dacă nu intervin priorităţi şi judecăţi de valoare care să filtreze aceste decizii. Încrederea oarbă în date şi în obiectivitatea sau imparţialitatea tehnologiei poate ascunde politica implicită a acestor modele şi poate îngreuna dezbaterea politică şi morală privind acţiunile viitoare. Mai mult, Big Data sunt, prin definiţie, acontextuale, cumulând fluxuri informaţionale derivate din situaţii de viaţă foarte diferite. În ce măsură pot aceste date să fie temeiul unei decizii ancorate în contexte bine înţelese, în detaliile particulare ale unei situaţii?

Nu în ultimul rând, asistăm la *degradarea cunoașterii colective și a capacității comunitare de rezolvare a problemelor societății*, prin răspândirea bulelor de conținut și a știrilor false. Big Data sunt asociate cu modelele de afaceri ale giganților tehnologici care oferă personalizarea

conținutului informativ în funcție de preferințele individuale ale consumatorilor. Asta duce la segmentarea fluxurilor de știri și la spargerea dezbaterii publice, atât de-a lungul unor falii majore, cât și în nenumărate bule de interese punctuale. În ce măsură mai este posibilă o dezbatere publică, dacă dispare acordul asupra a ceea ce este real și ceea ce este fictiv?

Toate aceste riscuri sunt cu atât mai relevante când se aplică justiției înseși. Sistemul judiciar a început să folosească tot mai frecvent soluții bazate pe algoritmi predictivi și pe inteligență artificială. Sistemele automate și de IA promit să rezolve probleme care țin de supra-încărcarea cu dosare a sistemului juridic sau de prejudecățile și părtinirile judecătorilor umani. Astfel de tehnologii sunt folosite pentru a estima riscul de delincvență sau riscul de recidivă, influențând felul în care justiția îi vede pe cetățeni ca fiind „periculoși“ sau nu.

În zona ordinii publice, vorbim, de exemplu, despre poliție predictivă (*predictive policing*), în care forțele de ordine sunt direcționate prioritar spre zonele calculate ca având un risc ridicat. În tribunale, judecătorii primesc scoruri de risc ale inculpaților, pentru a-i ajuta să decidă diferite elemente ale desfășurării procesului și ale sentinței. Felul în care sunt calculate aceste scoruri este adesea greu de înțeles, deoarece nu avem acces la metodologiile lor de calcul, tehnologiile fiind apărate de drepturi de proprietate intelectuală. Mai mult, deciziile bazate pe învățare automată sunt opace chiar și pentru cei care au acces la cod, datorită straturilor multiple de complexitate atât în program, cât și în date.

Prin urmare, idealul *transparenței* deciziilor juridice este pus în pericol. Aleš Završnik observă, în analiza sa[5], că utilizarea inteligenței artificiale în justiție și poliție are potențialul de a afecta drepturi fundamentale, precum *prezumția de nevinovăție*, dreptul la un *proces corect* (inclusiv prin interogarea încrucișată a martorilor), dreptul la un *tribunal imparțial*, principiul *non-discriminării* și principiul legalității bazat pe *standarde clare de dovedire a vinovăției*.

Într-un studiu privind dezbaterile referitoare la implicarea inteligenței artificiale în justiția din SUA, judecătorul Claudiu Drăgușin remarcă *opacitatea*[6] acestor tehnologii, care este un obstacol considerabil pentru un

sistem de justiție corect. Nevoia de transparență poate fi, parțial, asigurată prin algoritmi open-source și prin auditarea umană continuă a acestor decizii.[7] Totuși, complexitatea modelelor IA riscă să transfere decizii juridice spre cutii negre ale căror acțiuni nu sunt, cel puțin deocamdată, explicabile. Cum putem evalua legalitatea, moralitatea sau înțelepciunea unei decizii care nu poate fi înțeleasă de oameni?

În concluzie, nu doar erorile sau lipsa cunoașterii pot fi problematice etic atunci când sunt temei al acțiunii, ci și excesul de precizie, de persistență sau de informații ilegitime. Cunoașterea poate fi prea multă și prin raportare la costurile sale, deoarece dreptatea socială presupune și un echilibru al cheltuielilor și beneficiilor. Transformarea ordinii sociale și economice pentru a deriva valoare din Big Data și din avansurile ML/AI trebuie să conducă la adaptări ale ordinii sociale, în vederea apărării drepturilor și libertăților fundamentale ale persoanelor într-o societate tot mai mult mediată tehnologic și a împărțirii corecte a acestei valori.

După cum observă Dan McQuillan, expert în inteligența artificială și lector la Goldsmiths, Universitatea din Londra, modelele ML/AI au anumite afinități politice, derivate din proprietățile organizațiilor și statelor care le construiesc și le utilizează.[8] Sistemele IA sunt adesea lăudate pentru capacitatea lor de a reduce costurile în sistemele supra-încărcate ale serviciilor publice. Totuși, bugetele subdimensionate ale serviciilor de sănătate, transport sau educație nu sunt fenomene naturale, ci alegeri sociale. Supra-încărcarea acestor servicii sociale este o consecință a deciziilor luate la nivelul colectivității; ea derivă din politicile neoliberale, care privilegiază interesele capitalului față de cele ale cetățenilor sau angajaților. Odată inserate în administrații sau firme orientate spre maximizarea profiturilor și reducerea costurilor cu (aproape) orice preț, inteligențele artificiale vor servi la mecanizarea acestor procese și la împiedicarea înțelegerii consecințelor lor asupra publicului. Deciziile ML/AI sunt opace și imposibil de inspectat în detaliu, poleind cu iluzia obiectivității și preciziei un set de calcule cu un anumit grad de arbitrar și eroare. Prin reducerea sistematică a inteligibilității deciziilor, dar și a numărului de oameni implicați în procesele decizionale, sistemele ML/AI devin efectiv soluții de prevenire a reflexivității, observă McQuillan.

Pentru a ne adapta la noua ordine socială a capitalismului datelor, trebuie să regândim la nivel colectiv principiile dreptăţii sociale şi formele lor practice de implementare. Tot ce ține de adoptarea inteligenței artificiale este la alegerea noastră, pornind de la *unde*, *când*, *cât de mult*, *cum*, *cu cine* și *cu ce preț*. Ca societate, putem alege să păstrăm sau să delegăm controlul uman asupra anumitor procese decizionale; putem alege să interzicem complet anumite forme de mecanizare a diagnozei și tratamentului cetățenilor sau să lucrăm la nuanțe. Premisa de la care pornim este că societatea datelor vaste și inteligenței artificiale nu este inevitabilă, ci, dimpotrivă, reprezintă produsul alegerilor noastre.

6. Regulile jocului: spre un model european al autonomiei umane

Ai apucat pisica de coadă, învârteşte-o să nu te zgârie.

În România şi pe alte meleaguri, liderii îşi pun speranţele în modelarea prin inteligenţă artificială şi culegerea de date vaste pentru a prezice şi a reduce incertitudinile viitoare. Să fie datele cele care ne ghidează, inevitabil, spre o viaţă mai bună?[14]

În capitolele anterioare, am discutat riscurile pe care le produce cunoaşterea bazată pe date vaste şi învăţare automată atât prin formele sale excesive, cât şi prin cele deficitare. Tehnologiile informaţiei schimbă relaţiile sociale şi raporturile de putere dintre oameni şi societăţi. Prin realocarea cunoaşterii, unii câştigă control şi putere de predicţie, iar alţii ajung să fie, cu sau fără voie, controlaţi şi predictibili. Se spune că tehnologiile sunt neutre, dar impactul lor depinde de oamenii care le folosesc şi de regulile jocului în care sunt inserate. Am văzut că, dincolo de intenţiile individuale, există două variabile sistemice ce influenţează puternic impactul tehnologiilor: şi anume modelele de afaceri prin care sunt puse în practică şi modelele sociale prin care sunt reglementate. Relaţiile de putere dintre producători, utilizatori şi terţe părţi nu sunt intrinseci tehnologiilor, ci depind în esenţă de relaţiile sociale care le guvernează. O tehnologie care ne ajută să vedem ce vrem, azi, poate fi cea care ne face să fim văzuţi fără voia noastră, mâine.

Deciziile algoritmice şi Big Data oferă o sursă enormă de valoare, dar şi o vulnerabilitate pe măsură. Devenim tot mai transparenţi prin datele colectate despre noi. Cum ne asigurăm că utilizarea acestora rămâne transparentă pe termen mediu şi lung? Trebuie să dezvoltăm legi adecvate şi mecanisme democratice de verificare. Altminteri, ne încredinţăm viitorul organizaţiilor care stăpânesc aceste date. *Vocea cetăţenilor va fi tradusă şi reformulată de Google, Facebook, Amazon, dar şi de Uber, Tinder sau competitorii lor, care devin infrastructuri ale alegerilor noastre.*

Trăim într-o lume compartimentată în bule ale rețelelor de socializare și ale fluxurilor de știri personalizate. În România, suntem mai degrabă axați pe crizele locale; ajungem să acceptăm necritic definiția Uniunii Europene ca arie săracă în inovare și antreprenoriat. Tehnologia ne apare ca o soluție salvatoare și de invidiat, disponibilă mai ales vecinului din SUA sau Asia. Totuși, tehnologia este o unealtă, pusă în slujba unor interese. Agendele politice și comerciale sunt cele care definesc impactul tehnologiei, la bine și la rău.

În dezbaterea globală, tehnologiile informației sunt simultan mitizate, ca răspuns pentru marile probleme ale omenirii, dar și înfierate, ca începutul sfârșitului democrației liberale. Spațiul digital a devenit însoțitorul fidel al vieții fizice. Găsim acolo un rezervor de gânduri, relații și oportunități pe care le explorăm și le exploatăm în fiecare zi. Suntem atât creatori, cât și consumatori ai lumii digitale. Mai mult, suntem și consumați, prin *cookies* și alte firimituri digitale pe care le lăsăm în urma noastră, ca niște veritabili Hänsel și Gretel. În loc de vrăjitoare, ajungem să fim înghițiți de algoritmii predictivi ai firmelor, statelor și organizațiilor care vor să ne cunoască și să profite de această cunoaștere. În societatea de azi, bazată pe date vaste, învățare automată și inteligență artificială (Big Data, ML și AI), împărțeala cunoașterii și a puterii se schimbă de la o zi la alta, iar cine-mparte parte-și face.

Putem examina aceste noi inegalități pe două niveluri. În primul rând, apar modificări ale relațiilor dintre oameni și corporații, mai ales între utilizatori și platformele tehnologice. În al doilea rând, se schimbă echilibrul de putere între marii jucători din arena globală, mai ales între Uniunea Europeană, Statele Unite și China.

În fiecare zi, miliarde de utilizatori se bucură de beneficiile gratuite ale motoarelor de căutare, ale rețelelor de socializare și ale aplicațiilor digitale descărcate pe telefonul mobil. În pandemie, spațiul digital a oferit un colac de salvare pentru toate interacțiunile blocate în spațiul fizic, iar giganții tehnologici sunt printre marii câștigători ai acestei dezordini istorice, consolidându-și și mai mult statutul. Desigur, firmele sau partidele politice au încercat dintotdeauna să influențeze consumatorii și votanții. Astăzi însă platformele digitale globale, precum Google sau Facebook, le pun la

îndemână soluții de profilare și clasificare fină a indivizilor. Devine posibilă livrarea și primirea unor mesaje personalizate în orice micromoment din zi și din noapte, dat fiind că avem telefoanele mereu lângă noi. *Precizia*, *adaptarea în timp real* și *omniprezența* sunt cele care fac diferența între campaniile pe panouri publicitare și campaniile pe mobil. Ce rămâne oare din libertatea de gândire și de decizie a indivizilor, supuși unui bombardament fin în cele mai vulnerabile momente personale?

Regulile jocului nu sunt aceleași peste tot. SUA, Uniunea Europeană și China sunt exemple clar diferite de abordare a potențialului Big Data, ML și AI. În China, aceste instrumente sunt folosite pe tot spectrul influenței și controlului comercial și politic, fără o dezbatere democratică. Acest continuum merge de la dezvoltarea propriului internet și propriilor platforme digitale, apărate de The Great Firewall of China[1], până la supravegherea cetățenilor prin scorul de credit social[2] și la identificarea etniei acestora prin recunoaștere facială. Deși guvernele și companiile de pretutindeni sunt foarte precaute în a manifesta rezerve față de stilul autoritar al dezvoltării tehnologice în China, detectarea automată a etniei uigure a fost un prag în care controversa academică a devenit un semnal global de alarmă[3]. Soluțiile ML/AI dezvoltate în China pot avea un impact mult dincolo de granițele acestui stat, schimbând natura cetățeniei în toată lumea.

Până acum, SUA au avut orientarea cea mai favorabilă mediului de afaceri, mai ales marilor corporații. Această țară a sprijinit inovațiile disruptive de tip Silicon Valley și apariția unicornilor și a platformelor care au cucerit lumea. În acest proces, adesea, a ignorat drepturile digitale ale indivizilor sau micilor afaceri. Investigațiile recente ale Congresului SUA privind Big Tech sunt un semnal al unei posibile reacții, dar rămâne de văzut dacă vor conduce la o echilibrare a asimetriilor de putere.

Uniunea Europeană este cea mai atentă și mai inovativă în ceea ce privește interesele cetățenilor săi privind automatizarea, apărându-le drepturile la viață privată și la protecția datelor personale, la un nivel nemaiîntâlnit în lume. Există premisa creării unui model european al capitalismului datelor. Automatizarea muncii prin ML/AI în sectoare tot mai diverse riscă să devină un tsunami, care însă nu va afecta pe toată

lumea la fel.

În cele din urmă, discuţia despre tehnologie devine o discuţie despre democraţia liberală. Provocarea noastră, ca europeni, este să evităm mitul tehnologiei salvatoare şi învestirea datelor cu încredere oarbă. Trebuie să regândim regulile jocului automatizării şi utilizării Big Data. Pentru a avea o şansă reală, Uniunea Europeană trebuie să continue să ţină pasul atât cu transformările tehnologice şi economice, cât şi cu provocările politice şi etice pe care acestea le ridică.

Etica şi drepturile omului în societatea algoritmică: filozofia morală şi modelul european

Nu-i greu să te sui în sus, îi greu să te ţii acolo.

Specificul modelului european constă în încercarea de a consolida drepturile fundamentale şi de a preveni eroziunea lor în capitalismul datelor. Acest demers porneşte de la redefinirea vătămărilor posibile, în noua societate informaţională. Voi porni în această secţiune de la Carta Drepturilor Fundamentale a Uniunii Europene şi GDPR. Ambele propun o redefinire radicală a limitelor omului sau, mai precis, a posibilităţilor de a atinge sau de a lovi o persoană. În vremurile trecute, daunele erau mai ales corporale sau materiale, prin atacuri fizice asupra corpului sau proprietăţilor noastre tangibile. În prezent, avem un corp informaţional, prin care putem fi manipulaţi sau răniţi. Dreptul la protecţia datelor personale şi dreptul la viaţă privată, stipulate în Cartă, recunosc riscurile tot mai mari aduse de supravegherea statală sau comercială, mai ales prin mijlocirea mediului digital. GDPR defineşte şi mai clar dreptul la protecţia datelor personale, introducând şi limite în controlul oamenilor de către boţi sau sisteme automate.

În aprilie 2021, Uniunea Europeană a lansat AI Act (Legea inteligenţei artificiale)[4], o propunere de reglementare a produselor bazate pe inteligenţă artificială care vor fi permise pe teritoriul său. AI Act reprezintă un set de reguli ce sunt, actualmente, în curs de negociere şi finalizare. O altă reglementare privind partajarea datelor de către actorii comerciali şi guvernamentali („EU Data Act“), este, la rândul său, în curs de elaborare.

Aceste inițiative europene devin repere pentru producătorii de tehnologii digitale. Ele influențează regulile jocului la nivel global, stabilind liniile directoare ale limitelor puterilor noastre (drepturi și obligații) ca utilizatori și cetățeni în capitalismul datelor.

Vom discuta în continuare filozofia morală din spatele abordărilor europene, care reprezintă un cadru legislativ de pionierat, cu reflectări directe în legislația din California, Brazilia și China, printre altele. Apoi, în secțiunea următoare, vom explora semnificația acestora pentru viața noastră de zi cu zi.

Pentru a putea evalua din punct de vedere etic utilizările curente ale Big Data și sistemele de decizii algoritmice, este util să ne raportăm la un sistem de referință din aria filozofiei morale, care să ne ofere reperele și regulile de judecată. Filozofia morală include teoretizări diverse ale acțiunii, care pot conduce la evaluări divergente, fie accentuând principiile morale ale acțiunii (în orientarea deontologică), fie rezultatele finale ale acțiunii (în orientarea consecinționistă sau utilitaristă), sau tipul de caracter moral din care rezultă deciziile și acțiunile (în etica virtuții). Capitolul de față își propune să prezinte succint câteva astfel de problematizări etice ale societății algoritmice, pornind de la studiul lui Richard Herschel și Virginiei Miori[5], și să detalieze apoi abordarea kantiană sau deontologică, pe care o consider cea mai productivă intelectual pentru înțelegerea implicațiilor morale ale datelor vaste.

Herschel și Miori[6] discută patru abordări posibile în analiza etică a sistemelor sociale animate de Big Data: etica deontologică sau kantianismul, contractul social, utilitarismul și etica virtuții. *Etica deontologică*, la care Immanuel Kant a adus o contribuție semnificativă, este centrată pe conceptul de demnitate a ființei umane și pe necesitatea tratării tuturor persoanelor ca scopuri în sine, nu ca mijloace – ceea ce Kant a denumit „imperativul categoric“. Prin urmare, avem datoria morală de a respecta autonomia celorlalți și demnitatea lor ca scop în sine, nu doar ca mijloc pentru alte valori. De aici pot fi derivate o serie de drepturi și libertăți umane fundamentale, care operaționalizează această valorizare a demnității și autonomiei omului.

Pentru Herschel și Miori, această perspectivă este foarte relevantă în

judecarea soluţiilor bazate pe Big Data, deoarece problemele pe care acestea le pun în ceea ce priveşte respectarea autonomiei indivizilor au devenit evidente şi mult discutate. De exemplu, autorii observă că o perspectivă kantiană ar însemna să ne punem întrebarea dacă fiecare om ar trebui să accepte sau nu o regulă generală care ar permite „partajarea informaţiilor despre fiecare individ cu sau fără permisiunea sa, indiferent dacă sunt corecte sau nu, complete sau nu, actuale sau învechite, informaţii care ar putea să fie folosite pentru a influenţa şi a descrie comportamentul şi interesele oamenilor cu sau fără acordul lor“[7]. Desigur, aceasta este o întrebare retorică şi răspunsul este clar negativ. Remarcăm aici proeminenţa pe care o au erorile (deficitul de cunoaştere) şi potenţialul de a prezice – şi deci de a manipula – acţiunile oamenilor (excesul de cunoaştere) specifice soluţiilor Big Data.

Perspectiva contractului social, bazată pe lucrări ale unor gânditori precum Thomas Hobbes, John Locke, Jean-Jacques Rousseau şi John Rawls, întemeiază judecăţile morale în aranjamentele sociale care creează legitimitatea anumitor acţiuni sau valori. Societăţi diferite vor acorda importanţă variabilă chiar şi valorilor fundamentale, precum viaţa, libertatea sau bunăstarea. Putem observa această variabilitate, observă Herschel şi Miori[8], şi în privinţa priorităţii acordate protejării dreptului la viaţă privată şi a datelor personale. Dacă Europa oferă un statut privilegiat drepturilor şi libertăţilor privind viaţa privată, Statele Unite ale Americii oferă mai puţină importanţă, în vreme ce China se poziţionează la celălalt pol al continuumului. Herschel şi Miori observă că această perspectivă poate oferi repere utile în evaluarea practicilor privind Big Data. Totuşi, cred că este nevoie să problematizăm legitimitatea acestor contracte, care nu pot fi considerate inerent justificate. Prin urmare, în continuare avem nevoie să apelăm la teorii morale deontologice, utilitariste sau ale virtuţii, pentru a evalua, din punct de vedere etic, un anumit cadru legislativ acceptat într-o societate.

Perspectiva utilitaristă, inspirată de operele lui Jeremy Bentham şi John Stuart Mill, pune accentul pe consecinţele acţiunilor noastre sau ale regulilor pe care acţiunile noastre sunt întemeiate. Herschel şi Miori[9] consideră că această teorie etică este greu de aplicat în cazul Big Data,

deoarece costurile și beneficiile sunt foarte diverse și distribuite inegal între actorii sociali multipli. Cum ar putea fi ele aduse la un numitor comun, pentru a determina balanța de beneficii pentru societate, în ansamblul său?

Nu în ultimul rând, *etica virtuții*, inspirată de gândirea lui Aristotel și Platon, definește acțiunea morală ca fiind aceea care izvorăște dintr-un caracter virtuos, ce urmărește o viață bună, orientată de valorile onestității, generozității și dreptății.[10] Herschel și Miori observă că este dificil să judecăm virtutea acțiunilor celor ce pun în practică produse și politici bazate pe Big Data. Cum am putea oare evalua dacă aceste acțiuni sunt bazate pe generozitate sau avariție, pe onestitate sau pe dorința de manipulare în folos propriu?

Mai mult, cred că o problemă suplimentară se referă la posibilitatea transferului judecăților bazate pe etica virtuții dinspre persoane naturale înspre persoane juridice. Soluțiile Big Data sunt create și puse în practică de organizații, nu de indivizi, acestea derivând din strategii organizaționale și modele de afaceri care sunt agreate și rafinate în raport cu acționari și diferite părți interesate. Cum putem evalua virtutea unui model de afaceri? Cred că o posibilitate este să ne uităm la orientarea temporală, și anume în ce măsură sunt avute în vedere doar consecințele și beneficiile pe termen scurt, sau sunt factorizate și consecințele pe termen mediu sau chiar lung. Această abordare încurajează o sustenabilitate mai mare și o mai mare deschidere înspre probleme de reputație și de relație cu societatea în care activează respectiva organizație. Totuși, judecarea organizațiilor prin perspectiva virtuții rămâne o problemă spinoasă.

În continuare, voi discuta drepturile relevante reglementate în spațiul european și consecințele lor privind produsele și politicile Big Data. Aceste drepturi au evoluat în timp, ca o reacție la transformarea cetățeniei odată cu dezvoltarea sferei informaționale și digitale. Observăm cum, pe măsură ce datele au devenit tot mai importante în funcționarea societății și a economiei, dreptul la viață privată nu a mai fost suficient pentru a apăra integritatea persoanei, atât în sfera offline, cât și în sfera informațională online. De aici a apărut nevoia dezvoltării unui drept dedicat protecției integrității informaționale a persoanei. Este vorba despre *dreptul la protejarea datelor cu caracter personal*, care a început să fie

instituționalizat prin noi autorități publice menite să îl asigure, precum cele dezvoltate de GDPR, sau de Legea privind confidențialitatea consumatorilor din California (California Consumer Privacy Act – CCPA).

Deși dreptul la protecția datelor personale pare să fie o derivație a dreptului la viață privată, cele două sunt, de fapt, distincte.[11] Din punct de vedere cronologic, observă Kokott și Sobotta în studiul lor comparativ din 2013, dreptul la viață privată a fost recunoscut primul, în Convenția Europeană a Drepturilor Omului[15] (în continuare abreviată prin Convenția), acceptată de cele 47 de state ale Consiliului Europei (inclusiv, printre altele, Elveția, Rusia și Turcia).

Implementată din 1953, Convenția specifică, la articolul 8: „Dreptul la respectarea vieții private și de familie. (1) Orice persoană are dreptul la respectarea vieții sale private și de familie, a domiciliului său și a corespondenței sale. (2) Nu este admis amestecul unei autorități publice în exercitarea acestui drept decât în măsura în care acesta este prevăzut de lege și constituie, într-o societate democratică, o măsură necesară pentru securitatea națională, siguranța publică, bunăstarea economică a țării, apărarea ordinii și prevenirea faptelor penale, protecția sănătății, a moralei, a drepturilor și a libertăților altora“. Arbitrul final este Curtea Europeană a Drepturilor Omului din Strasbourg.

În anul 2009, a intrat în vigoare Carta Drepturilor Fundamentale a Uniunii Europene[16] (în continuare abreviată prin Carta), care a fost elaborată de către Comisia Europeană, Parlamentul European și Consiliul Uniunii Europene. Carta recunoaște și ea dreptul la viață privată și de familie, în Articolul 7: „Respectarea vieții private și de familie. Orice persoană are dreptul la respectarea vieții private și de familie, a domiciliului și a secretului comunicațiilor“. Totuși, spre deosebire de Convenție, Carta adaugă un drept dedicat protecției datelor personale, în Articolul 8: „Protecția datelor cu caracter personal. (1) Orice persoană are dreptul la protecția datelor cu caracter personal care o privesc. (2) Asemenea date trebuie tratate în mod corect, în scopurile precizate și pe baza consimțământului persoanei interesate sau în temeiul unui alt motiv legitim prevăzut de lege. Orice persoană are dreptul de acces la datele colectate care o privesc, precum și dreptul de a obține rectificarea acestora. (3)

Respectarea acestor norme se supune controlului unei autorități independente“.

În jurisprudența Curții de la Strasbourg, datele personale sunt, de asemenea, protejate, în virtutea dreptului la viață personală. Ce diferență face, atunci, specificarea unui drept suplimentar și explicit privind datele personale? Această întrebare este cu atât mai relevantă, cu cât dreptul la protecția datelor personale a fost în continuare operaționalizat prin Regulamentul General privind Protecția Datelor (General Data Protection Regulation – GDPR)[17], intrat în vigoare în 2018[18]. GDPR este valabil atât pentru statele UE, cât și pentru statele din Aria Economică Europeană (European Economic Area – EEA), reprezentând o detaliere a acestui drept cu implementare directă în toate statele, fără nevoia de a fi adaptat prin legislație suplimentară.

GDPR reprezintă, în practică, o consolidare semnificativă a dreptului privind protecția datelor personale. Din punct de vedere istoric, doi pași preliminari importanți au fost reprezentați de dreptul de auto-determinare informațională, stipulat de Curtea Constituțională a Germaniei pentru a proteja demnitatea persoanelor, și de jurisprudența franceză, care a derivat protecția datelor personale din dreptul la libertate[12]. O prezentare a diferitelor țări participante la aceste acorduri legale este prezentă în figura 35.

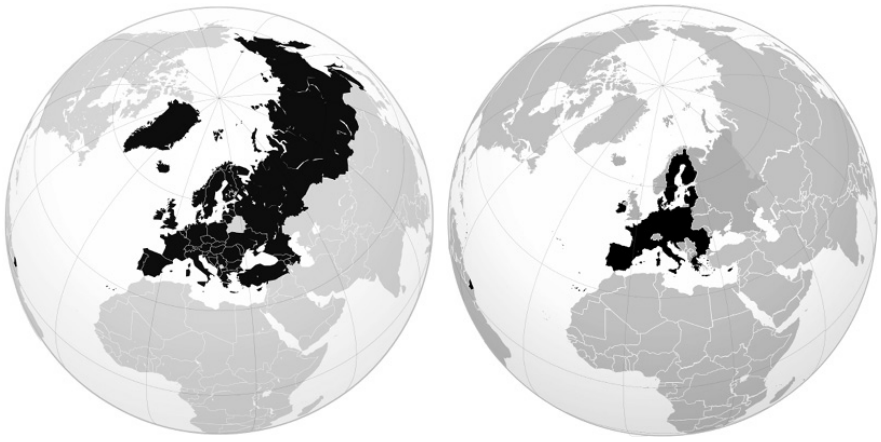


Fig. 35 Țările europene participante la acordurile legale relevante privind datele personale. Ilustrații preluate de pe Wikimedia (a) și de pe Wikipedia (b) și (c). **a** Țările Consiliului Europei, în care este acceptată Convenția Europeană a Drepturilor Omului. **b** Țările Uniunii Europene, în care este acceptată Carta Drepturilor Fundamentale a Uniunii Europene. **c** Țările UE și ale Ariei Economice Europene (EEA) în care este valabil regulamentul GDPR.

O primă distincție semnificativă apare în ceea ce privește aria de relevanță a celor două drepturi.[13] În primul rând, protecția datelor personale acoperă toate informațiile care se referă la persoane identificate sau identificabile, în timp ce protecția vieții private acoperă doar informațiile referitoare la viața privată a persoanelor. Deși viața privată este definită într-un sens larg, aceasta este totuși o îngustare a relevanței.[14] În al doilea rând, observă Kokott și Sobotta, Convenția se referă și la persoane fizice, și la persoane juridice, în timp ce Carta, prin implementarea sa în GDPR, acordă protecție doar persoanelor fizice.

O altă diferență importantă se referă la specificitatea dreptului și la distribuția responsabilităților. Carta a operaționalizat dreptul protecției datelor personale, indicând modalitățile acceptabile de colectare a acestora (consimțământ și motiv legitim prevăzut de lege). Totodată, a adus clarificări privind și două drepturi conexe ale subiectului datelor: dreptul de a-și accesa datele și cel de a obține rectificarea acestora.

De asemenea, este stipulat faptul că acest drept va fi asigurat prin activitatea unei autorităţi independente, nu doar prin infrastructura deja existentă de protejare a drepturilor omului. Mai mult, prin GDPR, *dreptul la protecția datelor personale a devenit nu doar responsabilitatea statului, ci și o responsabilitate a actorilor privați*, care trebuie să ia măsuri proactive de asigurare a respectării acestuia.[15] Se produce, astfel, o redistribuire a efortului, cu efecte sociale vizibile.

Această (re)alocare a responsabilităţilor şi a efortului de protejare a libertăţii individuale introduce o diferenţă crucială între dreptul la viaţă privată și dreptul la protecția datelor personale.[16] Rodotà observă că, pe de o parte, *dreptul la viață privată este conceptualizat mai ales ca un drept negativ – protejarea de interferența celorlalți – și un drept pasiv, care nu implică activităţi constante de menţinere, ci doar de reacție în caz de abuz.* Dimpotrivă, *dreptul la protecția datelor personale este un drept pozitiv și activ, implicând munca autorităților și a tuturor actorilor care procesează date, pentru a asigura respectarea constantă a acestei libertăți*.

Autorul observă că protecția oferită de GDPR este dinamică, urmărind fluxurile de date pe tot parcursul circulației lor între diferiți actori – ceva ce este clar mult prea complicat de realizat pentru o persoană. Autorități speciale sunt constituite și împuternicite să asigure protejarea datelor, spre deosebire de dreptul la viață privată care este asigurat de autoritățile statale generale. Aşadar, observă Rodotà[17], are loc o evoluție în complexitate de la dreptul de a fi lăsat în pace, respectiv conceptualizarea inițială a vieții private, la dreptul de a controla în mod activ construcția propriei sfere digitale și controlul informațiilor despre datele personale.

GDPR realizează o detaliere a dreptului la protecția datelor și, în acest proces, evidenţiază și valorile importante pe care acest drept se bazează. Aceste valori sunt esențiale pentru societatea europeană: protecția vieții private, transparența, autonomia și non-discriminarea.[18]

O consecinţă interesantă a explicitării dreptului la protecția datelor personale (prin Cartă și, ulterior, prin GDPR) constă în *extinderea simetrică a protecției acordate demnității oamenilor din spațiul fizic în spațiul digital.* Putem observa că, odată cu transferul unei părți tot mai mari din viața umană în spațiul digital, *persoanele fizice își extind aria de interacțiune și*

de vulnerabilităţi printr-o versiune digitală a lor.

Demnitatea şi libertatea persoanelor fizice depind tot mai mult de asigurarea unei protecţii active şi a unei libertăţi faţă de interferenţele şi manipularea acestei extensii digitale, prin care sunt mediate tot mai multe relaţii şi decizii semnificative. Rodotà[19] observă că dreptul la protejarea datelor personale reprezintă o recunoaştere a noilor realităţi derivate din inovarea tehnologică, care a generat un corp electronic al persoanelor, corp a cărui integritate trebuie să fie conceptualizată şi asigurată. Demnitatea omului constă, în secolul XXI, într-o integritate fizică, dar şi într-o integritate informaţională, corpul electronic dublând corpul fizic. Cetăţenia noului mileniu implică asigurarea spaţiului virtual în care oamenii să-şi poată exercita drepturile şi să-şi poată dezvolta personalitatea.[20]

Este important sa remarcăm faptul că protecţia datelor personale nu este limitată doar la spaţiul digital, ci acoperă orice tip de informaţii, în orice format, cât timp acestea intră sub incidenţa categoriei de date personale. Totuşi, discuţiile sunt cel mai adesea purtate cu privire la spaţiul digital, deoarece aici riscurile sunt mult mai mari, iar problematica abuzurilor datelor personale este legată de noile modele de afaceri ale giganţilor tehnologici. *Conceptul principal este însă mai degrabă cel al unei extensii informaţionale a persoanei fizice decât al unei extensii digitale* – cu observaţia că această extensie informaţională se petrece mai ales în mediul digital.

Prin urmare, putem observa cum *abordarea deontologică*, bazată pe recunoaşterea şi respectarea demnităţii şi autonomiei inerente fiinţei umane, *a permis, în ultimele decenii, o afirmare şi o consolidare a drepturilor fundamentale la protecţia vieţii private şi la protecţia datelor personale*. Abuzurile realizate în diferite episoade ale istoriei împotriva cetăţenilor – printre care al Doilea Război Mondial iese în evidenţă, în istoria recentă – au generat instituţii globale menite să protejeze drepturile omului, inclusiv în relaţie cu cercetătorii – aşa cum voi discuta în secţiunea următoare. Aceste drepturi devin tot mai actuale în contextul transformărilor economice şi sociale bazate pe Big Data, învăţare automată şi inteligenţă artificială.

De la vorbe la fapte: utilizatori, experţi, cetăţeni

Am eu ac de cojocul tău!

Tehnologiile care ne păzesc pot să fie, simultan, tehnologiile care ne rănesc, atât la nivel personal, cât și social. Digitalizarea, mai ales prin Big Data și inteligența artificială, au redefinit ce înseamnă cunoașterea, ce putem face cu ea și cine are acces la ea. Controversele privind influența acestor tehnologii asupra vieții noastre, de la psihicul copiilor la integritatea democrațiilor, sunt tot mai frecvente. Suntem oare condamnați să fim neîncetat spectatori ai scandalurilor și dezvăluirilor, sau putem transforma cunoașterea situației într-o strategie de răspuns?

Ce putem face depinde de ipostaza în care acționăm. Voi discuta mai jos câteva dintre aceste posibilități. De exemplu, probabil suntem, indiferent cât de ocazional sau de intens, *utilizatori* și *cumpărători* ai tehnologiilor digitale. Pentru acest rol au fost propuse numeroase strategii de acțiune și de preluare a controlului. Acestea merg de la refuzul unor tehnologii până la abordarea cui pe cui se scoate, la delegarea controlului către alți actori sau la gestiunea timpului petrecut pe ecrane. În al doilea rând, dacă suntem *specialiști* într-un anumit domeniu și înțelegem în profunzime impactul tehnologiilor digitale, putem acționa în interiorul organizației noastre, inclusiv în calitate de avertizori de integritate. Nu în ultimul rând, suntem *cetățeni* și, ca atare, putem apela la politicile publice care controlează aceste tehnologii. Această ipostază invită și la *acțiune colectivă*, prin susținerea unor organizații civice sau de jurnalism investigativ, implicate în monitorizarea tehnologiilor digitale. Fiecare dintre aceste ipostaze are propriile beneficii, dar și riscuri și limitări.

Utilizatori și cumpărători

Clientul nostru, stăpânul nostru?

În ziua de azi, în general, nu este potrivit să spunem că suntem clienții companiilor digitale, ci, mai degrabă, că suntem utilizatorii lor. Calitatea noastră de utilizatori–consumatori ne diferențiază de cea a firmelor care-și plasează reclamele pe ecranele noastre, ele fiind, de regulă, clienții plătitori.

Mai ales în cazul Google, Facebook și al altor platforme finanțate prin AdTech (vezi capitolul 2), atenția și intențiile noastre sunt vândute companiilor interesate. Totuși, denumirea de utilizator sau consumator este aplicată și atunci când noi înșine plătim pentru produse sau servicii. Ce implicații are această ipostază?

În 1997, Edward Tufte remarca[21]: „Există doar două industrii care își numesc clienții consumatori: cea a drogurilor și cea a computerelor". Profesorul american critica la acel moment presiunea comercială creată, prin marketing, de a cumpăra mereu versiuni noi de programe sau aparatură digitală, chiar dacă nu aduc un beneficiu față de versiunea anterioară. În 2020, când citatul său a fost preluat de *Social Dilemma*[22], pentru a ilustra nocivitatea social media, critica a fost redirecționată spre dependența psihologică pe care acestea o creează. Ipostaza de consumator sugerează vulnerabilitate, lipsă de putere, exact ca în cazul consumului de droguri. Suntem însă noi, utilizatorii, lipsiți de mijloace de apel? Cu alte cuvinte, suntem niște consumatori fără echivoc, prin excelență?

Pe măsură ce tot mai mult din viața noastră s-a mutat în mediul digital, s-a extins și suprafața de atac prin care devenim vulnerabili la intruziuni malițioase, la tranzacții înșelătoare sau la manipulări. Astfel, au apărut simultan tehnici preventive și instrumente defensive. De exemplu, *pentru sporirea securității cibernetice*, este util și chiar necesar, în anumite sisteme de operare, să folosim programe antivirus. „Paza bună trece primejdia rea", se spune. Dar ce înseamnă o pază bună? Ca și în mediul fizic, în care ne ferim de șarlatani sau de hoți, este nevoie și în spațiul digital să ne activăm un anumit nivel de vigilență, dar și să depunem un efort continuu de a ține pasul cu vremurile. Diferența constă, desigur, în viteza mult mai mare cu care se schimbă peisajul digital față de cel fizic.

Bunăoară, utilizatorii sunt sfătuiți să-și actualizeze continuu aplicațiile, pentru a reduce expunerea la vulnerabilitățile versiunilor trecute. Dacă explorăm zone sau scurtături periculoase, precum site-uri clandestine, aplicații din surse neverificate sau chiar și linkuri care par legitime, primite prin e-mail sau SMS, ne expunem unor atacuri. Ni se recomandă, așadar, să ne menținem doar în zonele sigure, verificabile, ale internetului. Prinși între tentația explorării și dorința de reducere a costurilor, uneori cu prețul

expunerii la majorarea – de altfel, nedorită și nerecomandată – a riscurilor, utilizatorii își găsesc, de regulă, o combinație proprie între precauție, vigilență și achiziția unor instrumente de protecție care le oferă măcar un minimum de siguranță și confort.

Similar acestor măsuri din zona securității, s-a dezvoltat treptat un repertoriu și pentru *reducerea expunerii datelor personale*. Avem de ales, adesea, între navigarea facilă și navigarea precaută și ușor împiedicată. Putem opta pentru viteză și confort, plătind cu datele noastre personale. Atunci putem da click rapid pe fiecare buton prin care suntem de acord cu utilizarea *all cookies*, ne autentificăm facil prin mediatori, precum Google sau Facebook, și rămânem mereu autentificați în toate aplicațiile noastre, cu un cont unic ce ne reprezintă și ne stochează tendințele și preferințele. Dacă, în plus, facem cumpărături cu carduri de fidelitate sau cu carduri de credit și avem mereu poziționarea geografică pornită, vom adăuga un strat consistent de date offline profilului nostru digital. Dacă, și mai mult decât atât, instalăm, ca niște veritabili exploratori digitali, toate aplicațiile care ni se par interesante, fiind automat de acord cu „Condițiile de utilizare“ (pe care nu le citim), sfârșim prin a deschide porți suplimentare brokerilor de date, prin care să ne urmărească fiecare gest.

Alternativele constau în acceptarea unor fricțiuni, restricții și eforturi suplimentare de control al scurgerilor de date. Ca utilizatori, putem alege să rămânem anonimi, să refuzăm cookies și să ne reducem amprenta digitală, folosind navigatoare și platforme care oferă protecție împotriva stabilirii unui profil. Putem investiga sistematic permisiunile pe care diferite aplicații sau site-uri le obțin, pentru a restrânge toate permisiunile excesive, dezinstalând aplicațiile care cer prea multe informații despre noi.

Astfel, există soluții dedicate de *monitorizare* și control a acestor scurgeri de date prin cookies, permisiuni automate și alte intruziuni ale companiilor tech în dispozitivele noastre. Putem alege și soluții de confuzionare a brokerilor de date (*obfuscation*), prin care să ne mascăm navigarea, creând zgomot. Astfel de programe fac click automat pe linkuri și reclame de care utilizatorul nu este interesat, în mod invizibil și discret, generând deci date false și contradictorii privind acțiunile persoanei profilate. Nu în ultimul rând, putem alege, pur și simplu, să nu (mai) folosim astfel de tehnologii

care ne supraveghează constant. Putem să apelăm doar la produsele cu o reputație bună în gestionarea datelor, dar și să ne limităm, măcar parțial, activitățile desfășurate prin tehnologii digitale sau care caută să ne alcătuiască profilul.

Dacă trecem de la ipostaza de utilizator spre ipostaza de cumpărător, observăm că uneori avem de ales între a plăti cu date și a plăti cu bani. Adesea, produsele sau serviciile digitale gratuite se susțin financiar prin reclame, care ne țintesc prin datele pe care le preiau de la noi. Dacă plătim costuri financiare suplimentare, putem încerca să ieșim din logica AdTech. Totuși, este important să observăm că această opțiune nu este întotdeauna clară sau automată. *Pe scurt, putem ajunge în situația în care să plătim și cu bani, și cu date.* Faptul că am cumpărat sau ne-am abonat la un serviciu digital nu înseamnă automat că firma nu colectează și nu vinde datele noastre către terțe părți. Dacă însă dorim să plătim doar prin consumul de reclame, probabilitatea este și mai mare ca noi să fim, de fapt, produsul care este vândut.

De exemplu, aplicații de dating, precum Tinder sau Grindr, oferă acces gratuit sau acces premium. Totuși, această opțiune nu schimbă tipul de date pe care compania îl culege despre noi sau tipul de acord pe care îl dăm pentru utilizarea datelor. Aplicațiile de dating au acces la un volum enorm de informații privind cele mai intime aspecte ale vieții noastre: de la orientarea sexuală până la poziționarea geografică. Numeroase investigații au arătat că aceste date sunt transferate spre brokerii de date[23] și pot ajunge astfel la terțe părți, agenți privați sau guvernamentali interesați de controlul acțiunilor noastre.

În iulie 2021, publicația catolică *The Pillar* a dezvăluit faptul că preotul Jeffrey Burrill, care deținea funcția de secretar general al Conferinței Episcopilor Catolici din Statele Unite, a vizitat baruri pentru homosexuali și diverse reședințe private. Aceste informații au fost obținute pornind de la datele despre locurile frecventate de preot, date stocate de aplicația Grindr. Cum a avut acces publicația la aceste informații? „Datele obținute și analizate de *The Pillar* includ semnale ale aplicației mobile pentru două intervale de 26 de săptămâni, primul în 2018 și al doilea în 2019 și 2020. Informațiile au fost obținute de la un furnizor de date și au fost autentificate

de o firmă independentă de consultanță, contractată de *The Pillar*.“[24]

Ca urmare a scandalului creat, preotul a demisionat din funcția de conducere în cadrul Conferinței Episcopilor Catolici. Impactul negativ asupra vieții sale personale și profesionale se va extinde, probabil, dincolo de pierderea acestui rol. Observăm cum disponibilitatea facilă a datelor digitale referitoare la acțiunile noastre poate dezvălui, cu ușurință, aspecte intime și potențial compromițătoare celor interesați și dispuși să plătească pentru acces. Reiese, de aici, o limitare structurală a rolului de utilizator sau de cumpărător. Această ipostază nu ne permite schimbarea semnificativă a regulilor jocului, ci doar o combinație selectivă de acceptare și refuz.

Voi menționa aici și posibilitatea unei alte ipostaze în care ne putem afla, și anume cea de *acționar*. Dacă ajungem să ne investim banii în portofolii de acțiuni, avem la dispoziție o nouă tactică de încurajare și susținere a acelor companii pe care le considerăm valoroase nu numai comercial, ci și etic. O evaluare a unei companii sau al unui fond de investiții pe metricile ESG – Environmental, Social and Governance (Mediu, Societate și Guvernanță) – reprezintă un posibil punct de plecare. Acești indicatori țin cont de măsurile pe care o companie le ia pentru a-și reduce amprenta asupra mediului și pentru a acționa responsabil și transparent în relația cu angajații, clienții și comunitatea. În septembrie 2020, Forumul Economic Mondial a propus o primă standardizare a acestor indicatori, care ar putea înlesni raportarea lor de către companii și ar face posibile judecăți comparative la scară largă[25].

Specialiști și profesioniști

Dacă ai intrat în horă, joacă!

Experiența unui loc de muncă într-o organizație ne schimbă adesea înțelegerea personală a responsabilității, libertății și constrângerii. În cadrul unei organizații, acțiunile noastre se combină cu ale altora, având efecte pentru mulți oameni. În același timp, ajungem să facem compromisuri între ce vrem și ce ni se cere, ce este fezabil sau ce ajută organizația să supraviețuiască.

Totodată, munca într-un colectiv ne oferă o putere mai mare decât

ipostaza de consumator individual. Putem, de exemplu, influența produsele și serviciile pe care organizația le oferă. Mai mult, ca profesioniști, câștigăm o înțelegere detaliată a provocărilor și posibilelor soluții și avem posibilitatea de a ne construi o rețea colegială care ne poate ajuta la nevoie.

De exemplu, ca utilizatori ai rețelelor de socializare, putem să ne gestionăm expunerea personală și, eventual, pe a celor apropiați. Totuși, nu vom avea puterea de a influența politicile Facebook sau Tinder și nici nu vom putea avea vreun impact semnificativ în viața multor alți oameni. Ca specialiști, dacă lucrăm pentru astfel de companii, vom înțelege mult mai bine ce se întâmplă. Vom putea participa la decizii și, dacă vom sesiza aspecte problematice, vom avea la îndemână măsuri pentru a le dezvălui, atât în interiorul, cât și în exteriorul organizației. Tot ca specialiști, vom putea crea noi produse care să ne ajute, pe noi și pe alți utilizatori, să ne apărăm drepturile și să ne trasăm obligațiile, în capitalismul datelor.

Principala diferență dintre ipostaza de specialist și cea de utilizator derivă din impactul final al acțiunilor noastre și din tipul de responsabilitate pe care o avem. Ca utilizatori, gestionăm riscuri pentru noi înșine și pentru cei foarte apropiați. Imprudențele noastre ne vor afecta direct și personal. Eventual, putem crea o problemă și pentru cei cărora le-am transmis un document infectat sau le-am recomandat o aplicație care recoltează masiv date.

Însă, ca membri ai unor organizații, deciziile noastre pot influența, chiar și indirect, viața a sute, mii sau milioane de utilizatori sau beneficiari ai produselor la care contribuim. Mizele finale sunt substanțial mai ridicate, dar responsabilitatea noastră individuală este diluată în colectivul organizației. Contează aici și poziția noastră în lanțul decizional. Cu cât puterea noastră de decizie este mai mare, cu atât impactul crește și responsabilitatea se accentuează.

Putem examina posibilitățile și limitările ipostazei de profesionist în lumina recentelor controverse legate de Facebook. O investigație din martie 2021 a *MIT Technology Review*[26] îl aduce în prim-plan pe Joaquin Quiñonero Candela, director de dezvoltare a inteligenței artificiale și coordonator al echipei Responsible AI din Facebook. Quiñonero a avut un rol esențial în crearea algoritmilor de personalizare a conținutului care

definesc modelul de afaceri și succesul comercial al Facebook. Aflat în prim-planul expansiunii companiei, Quiñonero este așadar legat atât de beneficiile pe care platforma le aduce utilizatorilor, cât și de impactul său nociv. Activitatea sa este direct conectată cu acuzațiile de manipulare electorală prin Cambridge Analytica, de facilitare a genocidului împotriva musulmanilor din Myanmar sau de polarizare socială și agravare a fenomenului *fake news*. Articolul lasă deschisă întrebarea: *Ce ar fi putut face Quiñonero pentru a remedia și a preveni astfel de probleme?* Sunt algoritmii săi o forță asupra cărora și-a pierdut controlul, sau mai are el încă puterea decizională de a redefini impactul Facebook? Este Mark Zuckerberg singurul om cu putere decizională în acest colos? Dacă nu, cine sunt ceilalți și cum participă la această responsabilitate colectivă?

Putem examina și o altă versiune a puterii specialiștilor. Timnit Gebru a ajuns în atenția publicului larg după ce a fost concediată de Google, din poziția de coordonatoare a echipei Ethical AI. Reacția a survenit într-un conflict pe marginea unui studiu privind riscurile modelelor lingvistice IA de mare anvergură publicat de Gebru împreună cu alții și intitulat „Despre pericolele papagalilor stocastici: pot modelele de limbaj să ajungă prea mari?“[27].

La vremea respectivă, studiul nu fusese încă prezentat propriu-zis mediului științific, fiind pus la dispoziția dezbaterii publice pe o platformă online, într-o etapă premergătoare publicării efective. Google i-a solicitat retragerea studiului, fără a-și detalia argumentele. Ulterior, compania a precizat că articolul nu ține cont de alte cercetări recente, ce ar fi îmbunătățit diagnoza. Gebru a refuzat să retragă articolul sau să își șteargă numele de pe publicație, fiind imediat concediată, conform declarațiilor sale. (Google a oferit o relatare alternativă, în care cercetătoarea și-a dat demisia din proprie inițiativă.)

Această măsură a fost considerată represivă de peste 1.200 de angajați ai Google și 1.500 de cercetători din mediul academic, care au semnat o scrisoare de protest.[28] Timnit Gebru își câștigase deja reputația în comunitatea științifică, prin documentarea părtinirii tehnologiilor IA. Printre altele, ea publicase un studiu împreună cu Joy Buolamwini, o altă cercetătoare de notorietate publică și academică. În acest studiu, cele două

autoare denunțaseră nivelul mai ridicat al erorilor de recunoaștere facială pentru femei de culoare decât pentru orice alt tip de respondenți, bărbații albi fiind cei mai avantajați.[29] Concedierea lui Gebru a fost urmată, în februarie 2021, și de concedierea colegei sale, coautoare a studiului incriminat de Google, Margaret Mitchell (semnatară a articolului final prin pseudonimul Shmargaret Shmitchell).

Vedem cum cele două specialiste au acționat în virtutea mandatului lor organizațional de a promova etica în domeniul IA. Ele au ajuns, chiar prin munca lor, într-un conflict puternic cu conducerea Google. Acțiunile lor au afectat reputația Google în comunitatea științifică[30], umbrind imaginea publică de forță a binelui pe care compania o promovează insistent. În ce măsură au influențat ele, direct sau indirect, și produsele Google bazate pe inteligență artificială, prezente sau viitoare? Despărțirea lor de companie nu sugerează un răspuns optimist.

Este important să remarcăm că atât Gebru, cât și Mitchell au o formație tehnică, în știința calculatoarelor. Preocupările lor privind discriminarea algoritmică și provocările etice ale IA vin, așadar, nu dintr-o poziție externă domeniului, ci chiar din înțelegerea nuanțată a acestor tehnologii, ele fiind direct implicate în crearea de algoritmi și modele.

O altă strategie disponibilă experților este cea a avertizorilor de integritate (*whistleblower*). Angajații care ajung să afle despre încălcări grave ale eticii în firma sau organizația lor pot divulga detaliile și documentele justificative jurnaliștilor, autorităților statului, site-urilor specializate în dezvăluiri sau publicului larg. Desigur, aceste acțiuni contravin obligațiilor contractuale și pot fi și ilegale, depinzând de protecția acordată acestui tip de protest în fiecare țară în parte. Prin urmare, avertizorii pot suferi consecințe grave profesionale, legale, dar și în viața personală, prin amenințări și hărțuiri.

Există cazuri celebre de *whistleblowing*, cu un impact social major. De exemplu, Jeffrey Wigand, prin dezvăluirile sale din anii ’90, a contribuit decisiv la schimbarea opiniei publice și a legislației privind fumatul. Din poziția de director adjunct la Brown & Williamson, o companie subsidiară a British American Tobacco, a avut acces la documente interne care demonstrau că țigările erau explicit produse pentru a spori dependența de nicotină.

În tot acest timp, managerii companiilor dădeau declarații conform cărora nicotina nu creează dependență. Wigand a expus publicului larg ipocrizia liderilor Big Tobacco și impactul catastrofic privind sănătatea. Prin colaborarea sa cu Food and Drug Administration (FDA) și prin informațiile oferite mass-media și mediului științific, Wigand a devenit o figură legendară a acțiunii etice îndreptate contra unor conspirații industriale împotriva binelui public.[31]

Companiile din domeniul IT nu au dus nici ele lipsă de avertizori de integritate. Într-un scandal celebru din 2015, mai mulți angajați ai startup-ului Theranos, printre care Tyler Shultz și Erika Cheung, au dezvăluit frauda în care era implicată compania, care pretindea că revoluționează analizele de sânge prin IT.[32]

În 2018, Brittany Kaiser și Christopher Wylie au dezvăluit[33] strategiile Cambridge Analytica de manipulare electorală. Alte companii, precum Pinterest, Amazon, Google sau Facebook, au ajuns în atenția publicului prin avertizorii de integritate, anonimi sau persoane publice, care protestau împotriva discriminării și a altor daune morale, sesizate din poziția lor de profesioniști.

În 2019, un angajat Google a dezvăluit detalii privind proiectul Nightingale, prin care gigantul tech a obținut zeci de milioane de informații medicale ale pacienților din 2.600 de spitale, gestionate de firma Ascension, al doilea cel mai mare furnizor de servicii de sănătate din SUA. Aceste informații includ date personale și istoricul medical ale indivizilor, care nu au avut nici un cuvânt de spus în transferul datelor medicale către Google.[34] Proiectul este în continuare în desfășurare, fiind însă și investigat de autorități.

Un alt caz celebru este cel al lui Frances Haugen, care, în 2021, a adus în atenția publicului și autorităților americane practicile comerciale ale Facebook prin care, consideră ea, compania pune profitul mai presus de interesul public (vezi și capitolul 1).[35]

Impactul avertizorilor de integritate depinde foarte mult de opinia publică și, mai ales, de acțiunile juridice sau legislative pe care le pot declanșa. Apare aici însă riscul fragmentării și individualizării protestelor. Avertizorii sunt foarte profitabili pentru jurnalele care publică investigațiile și,

eventual, pot câştiga un capital propriu de imagine. Este oare însă eroismul individual cheia unei reforme la nivel social? Vor fi actorii responsabili aduși în fața justiției? Pot dezvăluirile publice genera schimbări sistemice, dincolo de scandalul generat? Pe lângă aceste întrebări, care au o multitudine de posibile răspunsuri valabile în diverse domenii de interes în care acţionează avertizorii, îmi îndrept atenția spre domeniul tehnologiei, întrebându-mă: Oare vor avea zecile sau sutele de avertizori din Big Tech impactul pe care Jeffrey Wigand l-a avut asupra Big Tobacco? Dificultatea formulării unui răspuns indică atât analogia imperfectă între daunele produse de cele două industrii, cât și puterea diferită pe care o au asupra societăţilor în care funcţionează.

Eroismul individual al specialiștilor poate sfârși doar ca o notă de subsol în confruntarea cu interese comerciale organizate în platforme globale și companii cu o putere greu de estimat. În același timp, indignarea publică poate oferi resurse de legitimitate adoptării unor măsuri legislative majore, precum GDPR. Cum putem converti acţiunile morale individuale în schimbare socială? Cheia amplificării impactului protestelor individuale poate consta în energia colectivă deja mobilizată de organizaţii și asociaţii care luptă pentru apărarea demnității umane în fața eroziunilor tehnologice.[36] Ajungem așadar la a treia ipostază în care putem acționa, aceea de cetățeni și membri ai unei comunităţi.

Cetățeni și decidenți

Cine se aseamănă se adună.

O examinare a societăților prezente și trecute ne arată că nu există un răspuns unic la problemele traiului în comun. De exemplu, societățile sunt, într-o oarecare măsură, inegale. Totuși, în prezent ca și în trecut, unele societăţi tolerează sau chiar încurajează inegalităţi, care ajung până la dezumanizarea anumitor grupuri. Cât de diferite sunt standardele de viață ale oamenilor din straturile sociale de la baza ierarhiei față de cele ale indivizilor aflaţi la vârf? Cum sunt produse, impuse și justificate aceste diferențe? Există o varietate de modele și de opțiuni pe care le avem la nivel social, după cum am discutat și în capitolul 1. Aceste modele definesc

regulile jocului, dar rămâne și la latitudinea noastră ce fel de strategie de joc să adoptăm.

De exemplu, imediat după intrarea în aplicare a Regulamentului General de Protecție a Datelor (GDPR), asociația franceză La Quadrature du Net[19], alături de 10.000 de cetățeni, a dat în judecată compania Amazon, invocând faptul că folosește datele personale ale utilizatorilor pentru a-i ținti cu reclame personalizate, fără avea consimțământul lor liber exprimat.[37] Tot atunci, la Quadrature du Net, împreună cu 12.000 de cetățeni, a dat în judecată și Google, pentru aceeași impunere a reclamelor personalizate, fără obținerea acordului exprimat liber și transparent.[38] Ambele procese s-au soldat cu un verdict în favoarea utilizatorilor. Google a fost amendat cu 50 de milioane de euro, pentru lipsa de transparență în obținerea consensului de prelucrare a datelor când sunt create conturi noi. Amazon a fost amendat cu 746 de milioane de euro – cea mai mare sancțiune aplicată până în 2021 pentru încălcarea GDPR – pentru personalizarea reclamelor, folosind date obținute fără consimțământ. Maximilian Schrems, cetățean austriac și fondator al asociației NOYB.eu[20], a demonstrat, în instanță, că datele sale personale sunt transferate ilegal în Statele Unite, unde pot fi inspectate de serviciile secrete americane, încălcându-i drepturile de cetățean european.[39] Ca urmare a acțiunilor sale în justiție, acordul Privacy Shield privind transferul datelor din UE în SUA a fost declarat ilegal. Ca urmare a acestei decizii, în 2021, două instanțe diferite au ajuns la concluzia că Google Analytics nu respectă legislația europeană și drepturile cetățenilor europeni atunci când le transferă datele în SUA.[40]

Fiecare dintre noi exercităm putere asupra altora, suntem supuși puterii celorlalți și ne negociem preferințele și interesele în relații mai mult sau mai puțin echilibrate. În funcție de societatea în care trăim și de poziția pe care o ocupăm în ierarhiile sale, vom câștiga sau vom pierde grade de libertate în controlul asupra propriei vieți și al altor oameni. Noile tehnologii digitale și modelele de afaceri care le propulsează creează noi teritorii și deschid noi confruntări. Din cauza schimbării rapide a mediului digital, strategiile jocului social și economic se schimbă accelerat și impredictibil[41], iar regulile rămân într-un decalaj față de jocul real. Exemplele de mai sus ne arată însă că, mai ales prin acțiune colectivă, în spațiul european, putem

influența regulile jocului și putem provoca giganții tehnologici să-și schimbe practicile de stabilire a profilului și de supraveghere digitală a indivizilor.

Așa cum nu există sport fără reguli sau meci fără arbitri, o piață liberă depinde de respectarea normelor care previn trișarea și abuzul de putere. În capitalismul datelor, noile tehnologii produc forme noi de cunoaștere, care schimbă relațiile dintre jucători în feluri complicate și greu de anticipat. Ce impact va avea asupra muncii noastre și a generațiilor următoare automatizarea ocupațiilor prin IA? Cum vom fi percepuți de companii, state și terțe părți prin supravegherea continuă a acțiunilor noastre online și offline? Cum ne vom spori propria cunoaștere și putere prin conectarea la platformele digitale care cuprind și modelează societatea? Care vor fi raporturile cost–beneficiu pentru noi, pentru alți oameni ca noi și pentru cei ce trăiesc în alte straturi ale societății? Nu există răspunsuri evidente la aceste întrebări. Viitorul depinde de alegerile pe care acum le facem la nivel colectiv, în calitate de cetățeni europeni, dar și de cetățeni ai planetei.

De exemplu, în octombrie 2021, Parlamentul European a adoptat o rezoluție care solicită Comisiei Europene interzicerea utilizării recunoașterii faciale în spațiile publice de către poliție și organizații publice sau private.[42] Aceeași rezoluție cere și interzicerea folosirii tehnicilor de prevenție infracțională bazate pe IA de și a unor produse private de recunoaștere facială (precum cele oferite de firma Clearview; vezi capitolul 3) pentru identificarea unor suspecți. Acest apel, deși nu are deocamdată un statut obligatoriu, este un pas spre limitarea dezechilibrelor de putere pe care tehnologiile digitale le creează în capitalismul datelor. Nu este clar însă care va fi răspunsul legal ales de Uniunea Europeană.

Recunoașterea facială este un exemplu de teritoriu controversat, pe care vedem atât mișcări de avans, cât și mișcări de retragere. La finalul anului 2021, Facebook a anunțat că renunță la a mai folosi recunoașterea facială pentru a eticheta automat fețele din fotografiile postate.[43] Amazon și Microsoft au anunțat în 2020 o pauză în programele de dezvoltare de produse ML/AI pentru forțele de ordine.[44] În același timp, administrații publice și servicii private de autentificare folosesc tot mai des această metodă. După cum reiese dintr-o investigație a revistei *Wired*, recunoașterea

facială devine o tehnologie polarizată, fiind simultan interzisă în unele state și normalizată în altele.[45] În noiembrie 2021, Internal Revenue Service, mai precis Fiscul din Statele Unite, a lansat o platformă federală online care folosește recunoașterea facială pentru autentificarea indivizilor.[46] Același sistem de autentificare, furnizat de compania ID.me este folosit și de alte autorități publice din SUA. Agenția de ocupare a forței de muncă din California a remarcat că acest sistem a blocat 350.000 de dosare frauduloase depuse pentru obținerea ajutorului financiar în timpul pandemiei Covid-19. Totuși, un raport de audit arată și că aproximativ 20% dintre solicitanții legitimi nu au reușit să se autentifice prin ID.me. Sistemele de recunoaștere facială dezavantajează mai ales persoanele de culoare, femeile, persoanele cu gen nonbinar și pe cei defavorizați economic. Observăm, în această instanță, un exemplu de compromis între anumite tipuri de eficiență și anumite tipuri de toleranță față de eșec și discriminare, în care anumiți actori câștigă, dar alții pierd.

În 2022, urmează să fie negociat și finalizat AI Act, prin care Comisia Europeană va decide liniile directoare ale modelului european de gestionare a automatizării digitale și inteligenței artificiale. Parlamentul European formulează rezoluții urmărind influențarea versiunii finale a acestui document esențial. Desigur, Parlamentul European nu este singurul actor implicat. Reprezentanții politici ai statelor membre urmăresc promovarea unor interese diverse, care se află în conflict.[47] Unele dintre state, de exemplu, susțin utilizarea recunoașterii faciale pentru controlul populațiilor, în timp ce altele se opun. Unele state doresc o reglementare largă a automatizării digitale, în timp ce altele preferă o țintire specifică a produselor IA, fără a reglementa alte tehnologii predictive. Majoritatea statelor se opun creării unui scor social de ierarhizare a cetățenilor pe baza datelor disponibile despre aceștia, dar nu este clar dacă firmele private vor avea voie să implementeze astfel de ierarhii. Cât de multă libertate for avea start-upurile să experimenteze cu soluții IA, inovând într-un domeniu riscant? AI Act va arbitra între interesele divergente ale țărilor, companiilor și ale altor actori sociali implicați.

GDPR a schimbat regulile jocului privind circulația datelor, introducând noi repere, arbitri și sancțiuni. Astfel, indivizii, firmele, corporațiile și

statele au început să țină cont de drepturile și interesele cetățenilor europeni.[48] Impactul legislației este, desigur, controversat – fiind acuzată atât că realizează prea puțin, cât și că încearcă prea mult.

Dacă luăm în considerare intervalul 2018–2020, GDPR a condus la un volum total de amenzi de 245,3 milioane de euro.[49] Totuși, în 2021, o amendă de 746 de milioane de euro dată către Amazon (contestată de companie) arată și că nivelul ar putea fi radical diferit.[50] În Statele Unite, Comisia Federală de Comerț (FTC) a amendat Facebook cu 5 miliarde de dolari, în 2019, pentru scurgerea de date din scandalul Cambridge Analytica.[51]

Implementarea GDPR este greoaie și deciziile suferă din cauza fragmentării agențiilor de control care operează la nivel statal. Mai mult, giganții tehnologici investesc eforturi și bugete enorme pentru a submina și a preveni reglementarea colectării datelor atât în Europa, cât și în SUA.[52] Limitarea legală a practicilor care stau la baza expansiunii și profitabilității Big Tech va fi, inevitabil, contestată și dificilă. Totuși, mizele sunt critice. Astfel de legi europene vizează evitarea unui (nou) feudalism digital, prin asigurarea justiției sociale și salvarea democrațiilor în fața inegalităților tot mai pronunțate. Rămâne de văzut în ce măsură vor avea impactul scontat.

Uniunea Europeană a sancționat giganții tehnologici nu numai pentru încălcarea drepturilor cetățenilor privind datele personale, ci și pentru abuzurile împotriva competitorilor. Din cauza poziției lor cvasimonopoliste, marile companii, precum Google, Apple, Facebook, Amazon și Microsoft, au la dispoziție mijloace de a-și dezavantaja rivalii, sugrumându-le accesul la piețe, platforme și utilizatori prin multiplele circuite de vizibilitate de care dispun. Până în prezent, politicile antitrust europene au condus la amenzi substanțiale aplicate acestor giganți, atât de autoritățile naționale din Italia, Franța și Germania[53], cât și la nivelul Comisiei Europene[54]. Totuși, este dificil de estimat impactul pe care amenzile îl au asupra logicii de afaceri a companiilor. O analiză Statista estimează că Google câștiga în aproximativ 6 zile suma necesară achitării amenzii de 2,4 miliarde euro administrată de Comisia Europeană în 2017.[55] Sunt aceste amenzi sancțiuni simbolice, sau au un impact în reformarea giganților tehnologici?

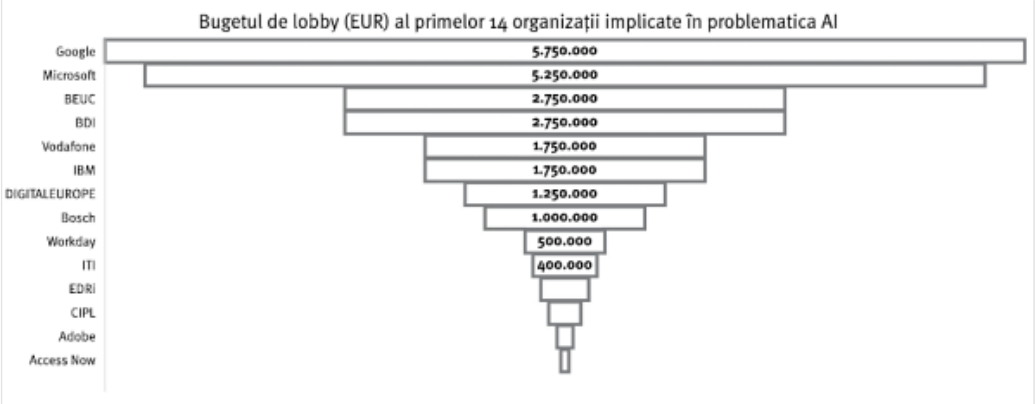


Fig. 36 Organizațiile comerciale domină lobby-ul privind reglementarea inteligenței artificiale în Uniunea Europeană, 2019–2021. Grafic realizat de autor pe baza datelor UE centralizate de EUobserver, 2021.

Marile corporații de tehnologie își mențin influența nu doar prin contestarea sancțiunilor juridice, ci și prin lobby, încercând să împingă produsele IA cât mai mult în organizarea viitoare a societăților și economiilor globale (vezi figura 36). În Uniunea Europeană, datele centralizate de EUobserver[56] privind lobby-ul arată bugete de milioane de dolari investite de Google și Microsoft pentru astfel de acțiuni.

În perioada 2019–2021, aceeași analiză arată că au fost înregistrate oficial 152 de întâlniri de lobby dedicate reglementărilor IA. Dintre acestea, două treimi au fost acordate companiilor comerciale. Marea majoritate a organizațiilor au participat la o singură întâlnire, dar Google se află în topul listei, cu 9 întâlniri (vezi figura 37). BEUC (Bureau Européen des Unions de Consommateurs), o uniune a 46 de asociații a consumatorilor din 32 de state europene, este cea mai vizibilă asociație nonprofit implicată în reglementarea IA.

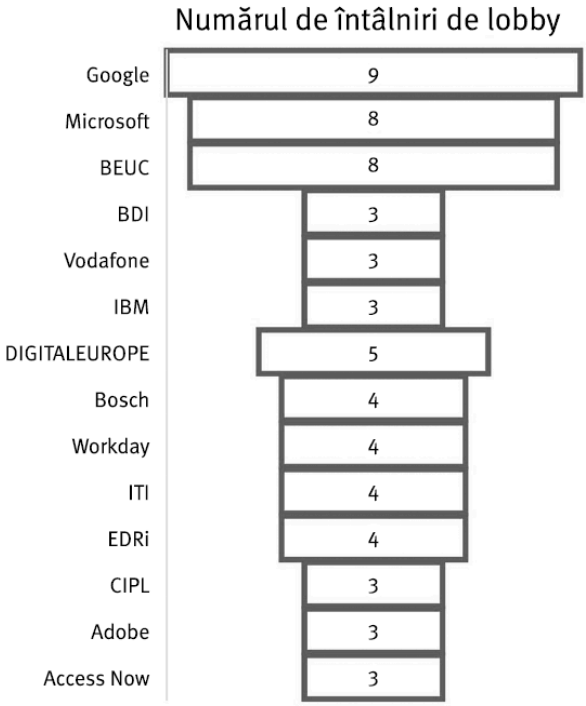


Fig. 37 Numărul de întâlniri de lobby ale principalelor organizații interesate de reglementarea AI, 2019–2021. Grafic realizat de autor pe baza datelor UE centralizate de EUobserver, 2021.

Influența Big Tech este și mai puternică în Statele Unite ale Americii. În perioada 2018–2021, managerii de vârf ai Oracle, Google, Microsoft și Amazon au colaborat în Comisia de Securitate Națională privind Inteligența Artificială, oferind recomandări președintelui și Congresului.[57] Aceste recomandări, incluse în bugetul din 2020, au propus măsuri concrete de achiziții ale unor soluții de IA în sistemul militar al SUA, precum și extinderea utilizării inteligenței artificiale în procesele administrative. O analiză *Wired* privind lobby-ul american referitor la tehnologiile de recunoaștere facială arată o intensificare considerabilă a presiunilor: numărul solicitărilor a crescut de patru ori din 2018 până în 2019, fiind în continuare în creștere în 2020.[58]

Revenind la viitorul AI Act, miza principală a eforturilor de influențare a legislației europene constă în puterea pe care producătorii produselor tehnologice o vor avea, în raport cu statele și cu cetățenii. Criticile aduse propunerii legislative se referă la numeroasele incertitudini privind evaluarea acestor tehnologii, odată ce sunt puse în aplicare. Contează foarte mult cine va clasifica nivelul de risc al tehnologiilor IA, cine va evalua

măsura în care au fost (sau nu) respectate precauţiile necesare prevenirii daunelor. De asemenea, vor trebui stabilite mecanisme pentru a asigura respectul necondiționat al drepturilor fundamentale, fără a le compromite într-un calcul pragmatic al beneficiilor comerciale și politice aduse de produsele IA. Implementarea efectivă a legii va depinde și de locul de organizare și gradul de fragmentare a arbitrajului. Marile companii promovează ideea autoreglementării, prin care ar ajunge să aibă o voce dominantă nu doar în proiectarea acestor tehnologii, ci și în evaluarea lor. O abordare centrată pe respectarea drepturilor și demnităţilor umane fundamentale ar acorda, dimpotrivă, un rol esențial instituțiilor publice și specializate ale justiției, urmărind independența de presiuni comerciale sau politice.

O comparaţie a modelului european de protecţie a datelor personale cu modelul american şi cu cel chinez este instructivă. Conform unei analize *The New York Times*, Statele Unite nu au încă o reglementare similară GDPR la nivel federal, ci doar în trei state: California, Virginia şi Colorado.[59] În majoritatea statelor din SUA, observă aceeaşi analiză, companiile pot vinde datele colectate de la utilizatori fără să-i avertizeze, brokerii de date le pot revinde și nu există nici obligația de a notifica eventualele scurgeri masive de date provocate de atacuri cibernetice. Prin urmare, puterea acordată companiilor de a colecta, tranzacționa și folosi legal date despre consumatori este, practic, nelimitată, fără o contrapondere în drepturi garantate indivizilor.

În același timp, Comisia Federală de Comerț a Statelor Unite (Federal Trade Commission, FTC) a sancționat prin amenzi considerabile eșecurile giganților tech de protejare a datelor cetățenilor[60], totalizând 5,8 miliarde de dolari – deci un alt ordin de mărime față de amenzile GDPR. Dintre acestea, cea mai mare amendă este de 5 miliarde de dolari, pentru Facebook, în scandalul Cambridge Analytica. Alte amenzi includ 575 de milioane de dolari plătite de Equifax pentru scurgerea de date din 2019, 170 de milioane de dolari plătite de YouTube pentru încălcarea protecției vieții private a copiilor și, tot în 2019, 5,7 milioane de dolari plătite de TikTok pentru colectarea ilegală a datelor minorilor.

Spre deosebire de SUA, China a optat pentru controlul strict al colectării

și circulației datelor pe teritoriul național, dar această legislație este relativ recentă. În septembrie 2021, a intrat în vigoare Legea protejării datelor personale, similară GDPR, prin care companiile trebuie să dea socoteală pentru informațiile pe care le colectează, fiind sever restricționat transferul datelor în afara granițelor Chinei. De asemenea, cetățenii pot solicita corporațiilor acces la datele lor, precum și rectificarea erorilor sau ștergerea acestora. Spre deosebire de GDPR, această lege oferă statului chinez acces privilegiat la datele cetățenilor, urmărind consolidarea puterii statului în raport cu corporațiile și cu populația. Abordarea Chinei este coerentă cu politicile sale de centralizare a puterii politice și de control social prin cenzura datelor, marcând un punct de turnură prin restrângerea libertăților anterior oferite corporațiilor IT. În același timp, statul chinez devine un broker de date cu acces la resurse imense, intrând astfel în competiție directă cu jucătorii globali din această piață. Astfel de volume de date masive, controlate aproape exclusiv de statul chinez, fără a fi împărtășite și cu alte state sau cu alți actori comerciali, îi vor oferi statului un avantaj considerabil în dezvoltarea produselor bazate pe învățare automată și inteligență artificială.

Observăm, așadar, confruntarea la nivel global a mai multe strategii de definire a regulilor jocului în capitalismul datelor. Prin implicarea noastră directă și prin sprijinul dat actorilor politici și civici care ne reprezintă interesele, precum asociațiile de protejare a vieții private și datelor personale pomenite mai sus (NOYB.eu și La Quadrature du Net, printre altele), putem contribui, în calitate de cetățeni, formatori de opinie și decidenți politici, la stabilizarea unui nou echilibru de putere și cunoaștere.

Mai mult, mai repede, mai opac: noi fronturi ale inteligenței artificiale

Pofta vine mâncând.

În iunie 2022, Blake Lemoine, un inginer implicat în comisia de etică IA a Google, a tras un semnal de alarmă care a ajuns în prim-planul atenției publice.[61] Lemoine a avertizat că LaMDA, tehnologia creată pentru a susține un dialog cât mai verosimil uman cu utilizatorii Google, are un nivel

de sofisticare atât de înalt, încât a căpătat conştiinţă. Prin urmare, a observat el, dorinţele, interesele şi drepturile LaMDA trebuie să fie respectate – deoarece, altminteri, Google creează o nouă formă de sclavie. Blake Lemoine a publicat şi un interviu cu LaMDA, realizat de el împreună cu un coleg, în care inteligența artificială răspundea la întrebări menite să-i sublinieze gândirea conştientă.[62] Deoarece superiorii ierarhici ai lui Lemoine au respins în repetate rânduri apelurile acestuia de a lua serios în considerare conştiința lui LaMDA, inginerul a început o investigație pe cont propriu. El a intervievat diferiți specialiști din afara companiei cu privire la implicațiile acestei situații și ce ar fi de făcut. Google a reacționat prin a-l suspenda din funcție, acuzându-l de încălcarea clauzelor de confidențialitate, ulterior fiind concediat.

Această controversă este semnificativă dincolo de miza apariției sau nu a unei forme de conștiință artificială și a eventualelor drepturi pe care ar trebui să le primească. Prin umanizarea LaMDA și prin alinierea cu dorințele atribuite acesteia, Lemoine ilustrează puterea ideii că IA poate deveni cvasiumană și poate avea gânduri și simțiri *aproape ca* oamenii. Oponenții acestei perspective argumentează că, indiferent de complexitate, inteligența artificială este, în cele din urmă, doar un prăjitor de pâine mai alambicat, un corector ortografic mai volubil sau un papagal stocastic.[63]

Nucleul chestiunii rămâne: *Cât de legitim este să încredințăm judecăți privind caracterul și competența oamenilor, precum și decizii privind soarta acestora, unor produse IA*? Vorbim despre admiterea la facultate, recrutarea pentru un loc de muncă, prioritatea pentru servicii medicale, accesibilitatea creditelor bancare sau a ofertelor imobiliare, dar și despre vinovăția sau nevinovăția în fața legii sau în fața unei drone de atac. Astfel de decizii ajung să fie tot mai mult luate de produse algoritmice masive, opace, adesea operate pentru profit și aflate în continuă schimbare. După cum observă Charlie Stross, riscul imediat al inteligențelor artificiale, pentru cei mai mulți dintre noi, constă în arbitrarul verdictelor automate, atunci când calculatorul spune nu, fără drept de apel.[64]

Aceste judecăți algoritmice colonizează zone tot mai ample ale vieții publice și private. Mai mult, expansiunea IA pare guvernată de o cursă spre fundul sacului. În această logică, dacă alții folosesc produse de inteligență

artificială, trebuie şi noi să facem la fel, ca să nu pierdem competiția. Trebuie musai să filtrăm tot mai repede volume de informații cu dimensiuni de nedescris. Trebuie să lucrăm cu mai puțini oameni, pentru a plăti mai puțin şi a obține mai mult de la cei cu care lucrăm. Trebuie să fim văzuți că folosim produse IA, pentru a părea o organizație inteligentă. Atractivitatea inteligențelor artificiale devine cu atât mai evidentă în contexte adversariale, când este imperativ să folosim orice avantaj pentru a domina inamicul.

Din păcate, viața socială nu duce lipsă de inimiciții, imaginare sau reale. Voi ilustra câteva procese în desfășurare de legitimare a folosirii inteligenței artificiale pentru a clasifica şi evalua oamenii, în trei domenii: în relațiile tot mai conflictuale dintre manageri şi angajați, în găsirea şi urmărirea în justiție a femeilor acuzate de avort, după criminalizarea acestui act în părți din SUA, şi în apărarea Ucrainei împotriva invaziei ruse.

Pandemia COVID-19 şi legitimarea IA prin managementul algoritmic

Cu cât te îmbogățești mai tare, d-aia te usuci mai mult.

Restricțiile impuse în pandemia COVID-19 au dus la extinderea muncii de acasă la niveluri de neimaginat înainte. Confuzia spațiului de muncă şi a spațiului privat şi suprapunerea temporalității de birou cu temporalitatea personală au stimulat eforturile şefilor de a supraveghea angajații. Au fost puse în funcțiune noi instrumente pentru a măsura cât mai precis timpul alocat muncii şi a sancționa abaterile de la disciplina organizației.

Digitalizarea a făcut posibilă supravegherea de la distanță cu un grad neașteptat de acoperire, dar şi de intruziune. Monitorizarea tastaturii şi a mouse-ului, capturile de ecran şi chiar fotografierea angajaților şi a mediului în care se află, adică a spațiului de acasă, au devenit tehnici standard şi foarte răspândite. Sistemele automate diagnostichează eficiența lucrătorilor pe baza acestor metrici. Apoi, îi sancționează prompt, atât financiar, cât şi prin îndemnuri de a-şi corecta conduita. Circulă povești distopice privind efectele distorsionante ale supravegherii, evaluării şi pedepsirii automate a angajaților. Investigația *New York Times* oferă un

repertoriu bogat de exemple.[65] Megan, care consiliază persoanele dependente de droguri, este pedepsită pentru timpul offline pierdut în discuțiile cu colegii despre cum să abordeze un caz dificil. Dacă nu tastezi, rezultă că nu muncești. Casierii din supermarketuri sunt sancționați, în numele eficienței, atunci când interacționează cu clienții vârstnici care se mișcă mai lent. Chiar și asistenții sociali care îngrijesc pacienți în faze terminale se plâng de nevoia de a acumula puncte de productivitate, alocate de un program care nu înțelege prea bine nuanțele fine ale finalului unei vieți omenești. Goana după puncte prin satisfacerea sistemului sau păcălirea acestuia devin motivații substanțiale ale angajaților nevoiți să obțină un scor bun pentru a-și câștiga salariul. Apare și o industrie contraofensivă. De exemplu, poți cumpăra un dispozitiv care să miște aleatoriu mouse-ul când vrei să mergi la baie sau să te desprinzi puțin de computer. Autorii investigației observă o inversare de rol: eram obișnuiți ca la serviciu să stăm cu ochii pe ceas, dar acum ceasul stă cu ochii pe noi.

Managementul algoritmic a fost inițial aplicat lucrătorilor care au slujbe precare și slab calificate, precum muncitorii din depozitele Amazon. Treptat, tehnologia a câștigat legitimitate și a început să guverneze angajați tot mai calificați, chiar și din poziții ierarhice ridicate. Cory Doctorow numește acest proces „curba adopției tehnologiilor mizerabile“ („shitty technology adoption curve“).[66] Cu decenii în urmă, observă el, doar prizonierii din închisori își luau cina sub camere de supraveghere. Azi, putem avea astfel de dispozitive în propria sufragerie, dacă suntem suficient de bogați și entuziaști tehnologic pentru sisteme smart-home. Același proces îl vedem și în cazul managementului algoritmic sau al produselor *bossware*, care au migrat din depozite și dispozitivele șoferilor Uber spre casele programatorilor, creatorilor de conținut și altor gulerele albe ce muncesc pe laptop pentru companii multinaționale faimoase. Carol, o expertă în domeniul finanțelor plătită cu 200 de dolari pe oră, constată, în investigația *New York Times*, că șefii săi o monitorizează și o remunerează în funcție de cât de intens tastează, în intervale de timp de zece minute. Minutele de reflecție sau scrisul cu pixul pe hârtie sunt sancționate scump.

Remarcăm în cazul tehnologiilor *bossware* atât excesul de precizie și de colectare intruzivă a informațiilor, cât și un evident deficit de înțelepciune.

Decalajul dintre cerinţele muncii din realitate şi imperativele de viteză şi de cantitate ale muncii aproximate prin programele de supraveghere conduc, adesea, la degradarea autonomiei angajaţilor, a complexităţii muncii şi a calităţii vieţii lor.

Cory Doctorow remarcă diferenţa esenţială dintre a folosi o tehnologie de inteligenţă artificială pentru a ne spori autocontrolul şi a o folosi pentru a-i controla pe alţii. Să punem în paranteză aici problemele legate de cuantificarea reducţionistă a vieţii, privind totul prin lentila metricilor de progres. Să punem în paranteză şi centralizarea datelor noastre personale de către firma care oferă tehnologia, scurgerile şi întrebuinţările licite şi ilicite ale acestora. Miliarde de oameni folosesc tehnologii digitale de automonitorizare, pentru a-şi înţelege şi a-şi îmbunătăţi rutinele şi abilităţile. Aşadar, o brăţară de fitness e foarte diferită de o brăţară electronică pentru controlul celor aflaţi în arest la domiciliu. Atunci când ne monitorizăm viteza de tastare pentru a ne eficientiza, e grozav; cu totul altceva este când salariul nostru depinde de numărul de clickuri pe minut. Utilizarea recunoaşterii faciale pentru a-mi porni telefonul mobil este foarte diferită de utilizarea sa pentru a-mi permite intrarea la metrou sau pentru a decide dacă sunt un pericol public în ochii guvernului.

O altă distincţie importantă apare între tehnologiile care clasifică şi prezic mediul fizic şi cele care evaluează persoanele şi anticipează deciziile umane. Cu rare excepţii, care ţin mai degrabă de impactul social indirect, folosirea inteligenţei artificiale pentru modelarea naturii nu stârneşte scandaluri publice. Nimeni nu se indignează din cauza predicţiei mai precise a vremii, a identificării structurii proteinelor, a detectării anomaliilor medicale în radiografii sau a anticipării nevoilor de reparaţii ale avioanelor. Cu totul alta este situaţia când inteligenţa artificială e pusă în joc pentru a modela şi influenţa alegerile noastre electorale, comerciale, romantice sau chiar pentru a ne reconstrui acţiunile şi motivaţiile şi a ne decide vinovăţia în faţa legii.

Criminalizarea avortului şi legitimarea IA prin supraveghere algoritmică

Cine stă în casă de sticlă, nu aruncă cu pietre.

În iunie 2022, Curtea Supremă a Statelor Unite a anulat decizia Roe vs. Wade, eliminând astfel dreptul constituțional la avort. Între timp, mai multe state din SUA au criminalizat întreruperea de sarcină, printre care Texas, Arizona, Oklahoma, Missouri, Arkansas, Louisiana, Mississippi, Alabama, Tennessee, Kentucky, Oregon și South Dakota.[67] Texas și Oklahoma permit cetățenilor să dea în judecată atât femeile care fac un avort, cât și persoanele care le ajută. Se deschide astfel calea spre acționarea în justiție și a celor care apelează la clinici din alte state, precum și a medicilor care efectuează avortul sau a celor care devin complici în ochii noii legi, ajutându-le pe femei în vreun fel. Criminalizarea avortului este intens controversată, stârnind emoții puternice de ambele părți. Prin urmare, energii considerabile sunt puse în mișcare atât pentru a prinde și pedepsi persoanele considerate vinovate, cât și pentru a le proteja, dinspre cealaltă parte. Datele masive și produsele de inteligență artificială vor deveni, cel mai probabil, resurse importante ale acestui conflict.

Activiștii antiavort, poliția și procurorii implicați în procese legate de avort folosiseră datele masive colectate online și anterior deciziei Curții Supreme[68], obținându-le prin mandat judecătoresc. Mai recent, investigații ale *Motherboard* au arătat că firme precum Placer.ai[69] sau SafeGraph[70] colectează și vând date privind poziționarea geografică celor ce vizitează centre de planificare familială, inclusiv clinici ce realizează avorturi. Deși datele sunt agregate, dezanonimizarea este posibilă, mai ales atunci când există puține cazuri într-o anumită zonă, ce pot fi identificate prin intersecția mai multor informații. Congresul SUA a început o investigație privind manipularea datelor legate de sănătatea reproductivă gestionate de brokerii de date, precum și de producătorii de aplicații specializate, de genul celor care monitorizează ciclul menstrual.[71] Între timp, femeile sunt sfătuite să anticipeze posibilele cumpărări sau obțineri prin mandat judecătoresc a informațiilor relevante. Este vorba despre indicii privind călătoriile lor, obținute din date privind localizarea geografică sau prin recunoașterea numărului de înmatriculare al mașinii, a istoricului căutărilor online și chiar a conținutului conversațiilor din rețelele de socializare, în eventualitatea în

care vor fi suspectate sau acuzate de avort. Localizarea geografică se numără printre cele mai sensibile informații și e cel mai greu de protejat. Practic, pentru a restricționa colectarea datelor privind deplasările noastre trebuie cel puțin să lăsăm telefonul mobil acasă, să evităm plățile cu cardul și să folosim transportul public.[72] Supravegherea video frecventă poate impune inclusiv acoperirea feței, ca măsură de precauție.

Dincolo de riscurile evidente ale urmăririi în justiție pentru o acțiune considerată acum ilegală, produsele de supraveghere algoritmică pun și alte probleme. De exemplu, ar fi posibil ca urme digitale ale intențiilor sau acțiunilor legate de avort să devină criterii discriminatorii în tot felul de decizii mediate digital. Algoritmii folosiți la angajare, la acordarea serviciilor sociale sau la evaluarea cererilor de cetățenie ar putea penaliza femeile considerate vinovate. Astfel de algoritmi sunt oricum opaci, adesea ascunși în spatele secretului comercial de către firmele care îi produc, fiind practic imposibil de inspectat de către marele public.[73] Discriminarea algoritmică pe baza deciziilor privind reproducerea poate fi deci realizată la scară largă, prin noile tehnologii bazate pe date masive și inteligență artificială. Chiar și date aparent inofensive din istoricul personal, precum căutările online legate de întreruperile de sarcină sau contracepție, pot expune persoanele la campanii de hărțuire și de demascare publică. Mai mult, acestea pot deveni dovezi în justiție, ca indicatori ale unor posibile intenții criminale.[74] Riscurile unor erori de detecție, a confuziilor sau interpretărilor greșite sunt enorme, putând duce la distrugerea vieții celor implicați în acuzații de avort. Controversa intensă privind criminalizarea întreruperii de sarcină ilustrează riscurile traiului în panopticonul digital și excesul intruziunilor pe care le înlesnește capitalismul supravegherii.

Ucraina și legitimarea IA prin armata algoritmică

Nu tot ce zboară se mănâncă.

Războiul din Ucraina este o altă arenă în care importanța inteligenței artificiale și datelor masive a devenit tot mai clară. Mai ales Ucraina și aliații săi, dar și Rusia, au pus în funcțiune arme autonome și alte instrumente bazate pe IA și pe analiza datelor vaste. Avântul automatizării

este susţinut de legitimitatea defensivă a Ucrainei şi de dorinţa statelor aliate de a o ajuta prin mijloace ce nu implică prezenţa personalului militar pe front. Prin urmare, remarcăm o sporire a legitimităţii utilizării inteligenţei artificiale în război. Putem observa această schimbare pe mai multe planuri, precum popularitatea enormă a dronelor, utilizarea identificării faciale sau dezvoltarea sistemelor IA de recunoaştere a terenului sau de mentenanţă predictivă.

Dronele au fost cea mai vizibilă dovadă a întrebuinţării inteligenţei artificiale în războiul din Ucraina, şi la propriu, şi la figurat. Într-un conflict puternic debalansat, utilizarea dronelor de partea Ucrainei a fost aclamată şi susţinută atât de guvernele ţărilor aliate, cât şi de populaţia acestora. Încă din februarie 2022, forţele ucrainene au folosit dronele turceşti Bayraktar pentru atacul convoaielor ruseşti[75]; în septembrie, preşedintele Zelenski anunţa construirea unei fabrici de astfel de vehicule autonome în Ucraina.[76] Pe lângă dronele turceşti şi cele autohtone, armata ucraineană a folosit şi drone fabricate în SUA[77], Norvegia[78], Polonia, Letonia şi China[79], în timp ce Rusia a utilizat atât drone din propriul arsenal, cât şi produse în Iran[80].

Războiul din Ucraina a stimulat o diversificare a utilizării dronelor atât pentru lansarea unor atacuri, cât şi pentru recunoaşterea teritoriului sau intimidarea inamicului. A apărut chiar şi un cuvânt nou ce surprinde sprijinul popular, şi anume „dronaţiile“.[81] Este vorba despre donaţiile în bani sau echipamente făcute, prin campanii de crowdfunding, de oameni din lumea întreagă pentru a susţine mai ales armata de drone ucraineană şi, în mai mică măsură, de partea cealaltă a frontului, pe cea rusă. De exemplu, campanii crowdfunding din Polonia, Lituania, Norvegia, Canada şi Ucraina însăşi au făcut posibilă achiziţia unor noi drone Bayraktar[82]. Trupa Kalush Orchestra şi-a vândut trofeul câştigat la Eurovision pentru a cumpăra, cu cei 900.000 de dolari obţinuţi, trei drone tip PD2, fabricate în Ucraina.[83] Democratizarea finanţării armelor autonome vine simultan cu o democratizare a echipamentelor, prin utilizarea celor de uz civil. Dacă o dronă Bayraktar costă aproximativ două milioane de dolari, una comercială DJI Mavic 3 costă circa două mii de dolari. Astfel, armatele actuale experimentează şi avansează continuu în utilizarea dronelor, conducând la o

normalizare a folosirii armelor autonome, atât pe front, cât și în imaginarul public.

Recunoașterea facială este o altă tehnologie a inteligenței artificiale folosită în războiul din Ucraina. Aceasta nu s-a bucurat însă de aceeași recepție călduroasă în mass-media. Platformele publice de recunoaștere facială au fost menționate încă de la începutul războiului, ca resurse în identificarea soldaților morți. De exemplu, FindClone estimează identitatea unei persoane folosind imaginile disponibile public pe internet, inclusiv pe rețelele de socializare.[84] Controversa s-a amplificat când, la începutul lunii martie 2022, Ministerul Apărării din Ucraina a anunțat parteneriatul încheiat cu Clearview AI.

Compania americană a oferit acces gratuit guvernului ucrainean la platforma sa de identificare facială.[85] Totuși, Clearview AI este foarte controversată, după cum am discutat și în capitolul 3, din cauza politicilor sale declarate de a aduna miliarde de fotografii postate pe rețelele de socializare, fără vreo permisiune a utilizatorilor sau companiilor vizate. Anterior războiului, protestele Twitter, YouTube, Google sau Facebook sau amenzile date de autoritățile din Italia, Regatul Unit, Franța și Australia nu au oprit această practică.[86] Guvernul SUA îi interzisese companiei să vândă accesul la această bază enormă de date companiilor comerciale, dar autoritățile publice americane continuau să aibă contracte cu Clearview AI.

Conform Reuters, în criza ucraineană, Clearview AI a anunțat și că are disponibile peste două miliarde de fotografii de pe rețeaua rusească VKontakte, care conține în total zece miliarde de imagini. Acestea permit identificarea mai precisă a soldaților ruși morți în Ucraina, în comparație cu platformele alternative. Guvernul ucrainean a folosit însă tehnologia Clearview AI și pentru a valida identitatea oamenilor vii, de exemplu la trecerea frontierei. Alte posibile utilizări ar include identificarea persoanelor pentru reunirea refugiaților cu familiile lor, identificarea celor suspectați de spionaj sau alte forme de susținere a armatei ruseşti sau demascarea știrilor false. Totuși, recunoașterea facială permite și numeroase utilizări ilegitime, așa cum a fost documentat din belșug prin investigații jurnalistice. De exemplu, cei care au acces la această tehnologie în virtutea postului pe care îl ocupă[87] pot folosi abuziv platforma și în scop personal – poate pentru a-și

urmări iubitul sau iubita, sau adversarii de un fel sau altul. Mai mult, guvernele pot apela la recunoașterea facială la scară largă pentru hărțuirea oponenților politici, inclusiv a jurnaliștilor, activiștilor ONG sau a judecătorilor. Rămân deschise și întrebările privind consecințele unei identificări eronate. Mai ales în timp de război, o etichetă greșită pusă pe fruntea cuiva poate fi diferența dintre viață și moarte.

În vreme de război, inteligența artificială apare și în tehnologii mai puțin captivante pentru marele public, mai ales pentru că nu se aplică direct oamenilor. Este vorba mai ales despre instrumente de *mentenanță predictivă*, tehnologii deja clasice în arii ale economiei civile. De exemplu, în industria aviatică, echipamentele de zbor conțin milioane de componente, multe dintre ele complexe și foarte specifice. Dacă firmele ar aștepta să se strice una pentru a comanda o piesă înlocuitoare, ar apărea întârzieri și costuri enorme. Modelarea echipamentelor și fabricilor ca sisteme supuse predicției prin inteligență artificială a permis intervenții corective anticipative, scăzând riscurile și reducând costurile. Războiul este o nouă arenă de utilizare a mentenanței predictive. De exemplu, experții militari americani dezvoltă metode de a anticipa nevoile armatei ucrainene privind armele și munițiile, înainte de a primi cereri explicite de aprovizionare. O astfel de tehnologie predictivă ar putea fi folosită și pentru analiza anomaliilor, răspunzând cerințelor de transparență cu privire la utilizarea resurselor oferite guvernului ucrainean.[88] Într-un alt scenariu de utilizare a inteligenței artificiale, UNDP dezvoltă o procedură de clasificare automată a daunelor și nevoilor de reconstrucție, mai ales privind infrastructura, prin analiza știrilor.[89] Inteligența artificială este folosită și pentru detectarea elementelor de interes în teren, precum trupe sau construcții inamice, pe baza imaginilor oferite de sateliții publici și privați, sau pur și simplu pentru prezicerea mai precisă a vremii, un factor important în orice confruntare pe teren.

Cu excepția recunoașterii faciale, care rămâne un subiect puternic controversat, *celelalte tehnologii ale inteligenței artificiale folosite de armata ucraineană și aliații săi par să fi preluat, prin contagiune emoțională, legitimitatea efortului defensiv al acesteia.* Observăm, în acest caz, ilustrarea pragmatică a dictonului „scopul scuză mijloacele“. Nu este

vorba doar despre discursul media. Sunt importante atât absența reglementărilor, cât și bugetele statelor aliate alocate acestor tehnologii. De exemplu, folosirea armelor autonome se extinde, fără mari speranțe de obținere a unor restricții la nivelul ONU privind roboții ucigași.[90] Tot mai multe țări investesc sume considerabile în tehnologii IA de uz militar, arată o analiză a MIT Technology Review.[91] În iunie 2022, NATO a anunțat un fond de un miliard de dolari pentru start-upuri din domenii precum inteligența artificială, datele masive și automatizarea. Regatul Unit a anunțat o strategie dedicată folosirii IA în domeniul militar, iar Germania se orientează și ea tot mai mult în această direcție, financiar și strategic.

În acest context, sporesc presiunile spre glorificarea inteligenței artificiale ca tehnologie a salvării umanității, produs de geniu ce nu trebuie să fie reglementat pentru a nu-i ofili creativitatea.[92] Contraponderea vine din partea unui discurs al riscurilor și incertitudinilor legate de impactul acestor tehnologii în lumea reală. Adesea, realitatea în care sunt efectiv implementate instrumentele IA adesea diferă semnificativ de condițiile în care au fost testate, cu atât mai mult în caz de război. Deciziile eronate sunt greu de identificat din cauza complexității și schimbării continue a tehnologiei. Alocarea responsabilității pentru greșeli este un coșmar. Promisiunile de implicare semnificativă a decidenților umani în operarea acestor tehnologii sunt adesea nerealiste și contrazise flagrant de funcționarea lor reală.[93] Rămâne de văzut cum poate fi ținută sub control această cursă tot mai accelerată, în plină desfășurare, cu energii financiare și mize militare și politice enorme.

Contraofensiva interesului public

Cine se frige cu ciorbă suflă și-n iaurt.

Asociațiile cetățenești, organizațiile neguvernamentale și guvernele înseși reacționează, în diverse feluri, la riscurile supravegherii algoritmice și abuzului puterii cvasimonopoliste a marilor platforme digitale. De exemplu, un nou proces de tip acțiune colectivă a fost intentat în California, SUA, împotriva Oracle. Compania este acuzată de acumularea unor informații excesive și intruzive referitoare la peste cinci miliarde de cetățeni, devenind

astfel o mașinărie globală de supraveghere care încalcă dreptul la viață privată a majorității locuitorilor planetei.[94] La rândul său, statul californian a intentat un proces antitrust companiei Amazon.[95] Procesul denunță practicile gigantului de a sili comercianții de pe platforma sa să vândă la prețuri cel puțin egale atunci când operează în alte piețe. Această cerință apare în condițiile în care, profitând de poziția sa dominantă, Amazon cere comisioane și taxe care ajung până la 35–45%, observă acuzatorii. Aceste politici extractive și coercitive au condus la o creștere generală a prețurilor pentru consumatorii din California, argumentează aceleași autorități. Creșterea profitului Amazon s-ar realiza, așadar, pe seama bunăstării cetățenilor.[96]

În SUA, inițiativele de reglementare rămân mai degrabă fragmentate și la nivel statal, nu federal. În Uniunea Europeană, inițiativele vin mai ales din partea Comisiei Europene și a Parlamentului European. Recent, Parlamentul UE a adoptat un pachet legislativ care definește un set de reguli privind piețele digitale. Legea Serviciilor Digitale și Legea Piețelor Digitale urmează să intre în aplicare în 2024, având scopul de a echilibra relațiile asimetrice de putere între platforme și utilizatorii acestora.[97] Măsurile stipulate urmăresc tocmai să prevină abuzurile monopoliste împotriva comercianților ce participă pe piețele agregatoare de tip Amazon, Google Play sau iStore, abuzuri anterior documentate de Comisia Europeană în procesele sale antitrust.[98]

De asemenea, rețelele de socializare vor avea propriile responsabilități de transparență și restricționare a daunelor. De exemplu, sunt interzise țintirea cu reclame personalizate a copiilor și a categoriilor definite etnic, politic sau prin alte trăsături protejate legal. Legislația europeană solicită platformelor, printre altele, să ofere utilizatorilor opțiunea reală de a nu li se alcătui un profil digital, să prezinte clar funcționarea algoritmilor de recomandare, să ofere comercianților instrumente pentru a-și valida pe cont propriu eficiența reclamelor, precum și să permită accesul cercetătorilor la datele necesare auditărilor independente.

Funcționarea GDPR este în continuare constrânsă de eficiența limitată a instituțiilor naționale responsabile cu punerea sa în funcțiune. Pentru a evita o gâtuire similară, noile legi europene prevăd înființarea unor autorități de

implementare specializate, la nivelul Comisiei Europene. Desigur, rămâne de văzut cum va funcţiona, în practică, această confruntare de termen lung între interesul public al cetăţenilor europeni şi interesele comerciale ale marilor companii de tehnologie.

De la societatea cunoașterii la capitalismul supravegherii

Ai grijă ce îți dorești, că s-ar putea să și primești.

În 1959, Peter F. Drucker formula conceptul de muncitor al cunoaşterii (knowledge worker) pentru a semnala importanţa tot mai mare a cunoaşterii în economia prezentului şi viitorului.[99] Mai mult, observa el, înţelegerea modernă a lumii în termeni de cauză şi efect este, treptat, înlocuită de o înţelegere în termeni de tipare. Apărea o nouă societate a cunoaşterii, dominată de cercetarea tiparelor din sistemele fizice, biologice, sociale, ecologice, cosmice etc.

Astăzi, şaizeci de ani mai târziu, societatea cunoaşterii (*knowledge society*) este un termen a cărui utilizare s-a plafonat. Nici măcar societatea informaţională (*information society*) nu mai pare să descrie satisfăcător lumea în care trăim. O examinare a popularităţii acestor sintagme ca termeni de căutare pe Google Search indică ascensiunea spectaculoasă a capitalismului supravegherii (*surveillance capitalism*) ca definiţie a prezentului (vezi figura 38). *Societatea cunoașterii* a devenit *societatea datelor*, îndreptându-se tot mai mult spre *societatea inteligenței artificiale*.

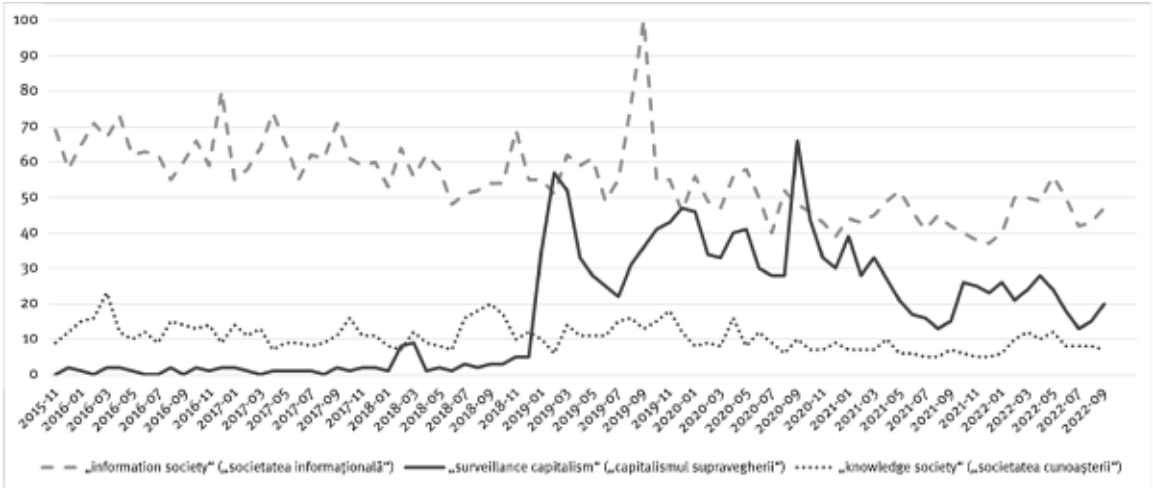


Fig. 38 Evoluția în timp a popularității sintagmelor „information society”, „surveillance capitalism” și „knowledge society” în termenii căutați pe Google, 2015–2021. Sursa: Google Trends, accesat pe 10 noiembrie 2021.

Specificul acestei tranziții constă nu doar în rolul tot mai mare al datelor ca resursă economică. Cele două trăsături definitorii sunt *renunțarea la înțelegere în favoarea predicției și delegarea deciziilor umane către sisteme automate.* Dacă muncitorii cunoașterii erau, în deceniile trecute, responsabili cu înțelegerea ordinii lumii și cu diagnoza sistemelor, astăzi *tiparele* sunt detectate prin învățare automată și nu mai presupun înțelegerea proceselor care le-au produs. Odată ce clasificările și ierarhiile trec din responsabilitatea oamenilor înspre cea a sistemelor digitale, deciziile devin și ele automatizate. Algoritmii ajung să evalueze, cu o implicare umană tot mai redusă, ce elemente sunt „cele mai bune“, „cele mai slabe“, „tipice“, „atipice“ sau „extraordinare“ – fie că este vorba despre oameni, animale, vehicule, echipamente sau orice alte *elemente* ale unor mulțimi finite sau infinite descrise de date.

Sigmund Freud observa că omenirea a trecut prin trei lovituri devastatoare ale percepției pe care o avea despre rolul său în lume.[100] Prima criză a apărut atunci când Copernic a arătat că Pământul nu este centrul universului. A doua criză, însoțită de o opoziție la fel de înverșunată, a survenit când Darwin și Wallace au arătat că oamenii nu sunt radical diferiți de alte specii. Freud considera că psihanaliza reprezintă o a treia lovitură de grație în amorul nostru propriu, dislocând individul din ipostaza de stăpân al minții sale. Capitalismul datelor și inteligenței artificiale promite o a patra detronare, care este dublă: oamenii pierd simultan pozițiile de singure *entități cunoscătoare și decidente*.

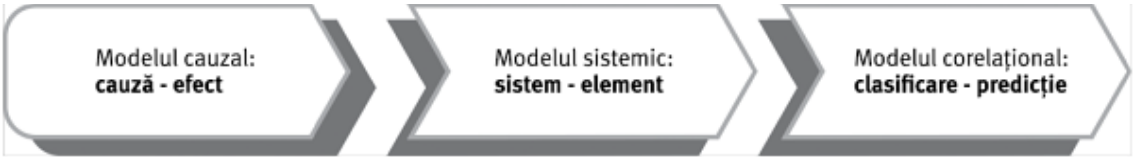


Fig. 39 Tranziția de la o înțelegere umană a cauzelor la o diagnoză umană a sistemelor și la o ordonare automată a elementelor. Ilustrație realizată de autor.

Inteligența artificială și analiza datelor vaste augmentează gândirea

umană. Miliarde de oameni folosesc produse digitale pentru explorarea unor lumi anterior inaccesibile. Memoria noastră colectivă, percepția spațiilor distante sau planificarea viitorului sunt azi profund influențate de aceste tehnologii. Trăim într-un preaplin al cunoașterii, în care provocările *selecției* sunt adesea mai mari decât provocările *accesului* la informații. Acest avânt informațional a fost convertit în bunăstare materială, în progrese științifice și medicale și în noi forme de expresie, creație și relații cu semenii. Câștigurile obținute ne înconjoară la tot pasul. Mai mult, bugetele de marketing ale companiilor reiterează continuu, în imaginarul public, promisiunile salvatoare ale tehnologiei. Cartea de față și-a propus să fie un moment de reflecție și pentru explorarea reversului medaliei, a excepțiilor, externalităților și efectelor perverse ale acestui preaplin.

Cele mai multe controverse iscate de inteligența artificială și de datele masive apar atunci când aceste forțe sunt îndreptate spre evaluarea, anticiparea și influențarea altor oameni. Capacitatea tehnologiei de a modela și controla natura, precum și abilitățile noastre de autocontrol câștigate prin aceleași metode riscă să devină și forțe ale exploatării, degradării autonomiei sau dezumanizării. La scara întregii planete, pe măsură ce comportamentul și modul de gândire ne este tot mai intens monitorizate, clasificate și prezise, frâiele deciziilor individuale și colective încep să fie preluate de sistemele automate și de organizațiile care le produc și le utilizează – fie ele corporații, guverne sau alte alianțe. Adesea, aceste sisteme sporesc prosperitatea colectivă, dar, alteori, creează sau reproduc vulnerabilități.

Cu 2.500 de ani în urmă, Heraclit observa că nu putem intra de două ori în apa aceluiași râu, pentru că nici noi, nici râul nu rămânem la fel. Astăzi, lumea este și mai schimbătoare. Nu putem călători de două ori cu același Uber, deoarece algoritmii de găsire a mașinii, de motivare a șoferului și de stabilire a traseului se pot schimba de la o cursă la alta. Inteligențele artificiale și datele pe care se bazează sunt atât de complexe și de dinamice, încât deciziile lor rămân esențialmente o cutie neagră, chiar și pentru cei care le creează. Așadar, într-o astfel de lume radical transformată prin tehnologie, auditarea și corectarea derapajelor devin tot mai dificile.

Cum putem deci negocia direcțiile în care ne conduc aceste inteligențe

artificiale construite pe principiul disruperii ordinii prezente? Ne regăsim în căutarea unui nou echilibru între complexitate şi bucle de reglaj, care să ne sporească bunăstarea, fără să ne vulnerabilizeze.

---

14. Această secţiune a fost publicată online pe platforma Republica.ro (R. Rughiniş, 22 aprilie 2021, „Salvarea prin tehnologie – o promisiune sau un avertisment? De ce să evităm mitul tehnologiei salvatoare şi investirea unei încrederi oarbe în date“, disponibil la adresa https://republica.ro/salvarea-prin-tehnologie-o-promisiune-sau-un-avertisment-de-ce-sa-evitam-mitul-tehnologiei-salvatoare).

15. Disponibilă online la: https://www.echr.coe.int/documents/convention_ron.pdf.

16. Disponibilă online la: https://eur-lex.europa.eu/legal-content/RO/ TXT/PDF/?uri=CELEX:12012P/TXT&from=PT.

17. Denumirea completă a GDPR este: „Regulamentul (UE) 2016/679 privind protecţia persoanelor fizice în ceea ce priveşte prelucrarea datelor cu caracter personal şi privind libera circulaţie a acestor date şi de abrogare a Directivei 95/46/CE“, disponibil în limba română la: https://www.dataprotection.ro/?page=noua%20_pagina_regulamentul_GDPR.

18. Disponibil online într-un format uşor accesibil la: https://gdpr-info.eu.

19. Prezentă online la https://www.laquadrature.net.

20. Prezentă online la https://noyb.eu/en.

Mulțumiri

Mulţumirile dintr-o carte sunt ca sarea în bucate, fiindcă nu pot lipsi şi nu sunt niciodată de ajuns, dar sunt esenţiale. Întrucât am multe cuvinte de recunoştinţă de împărţit, voi începe fără amânare. Le mulţumesc fiilor mei, Caius şi Remus, şi soţiei mele, Cosima, pentru curiozitatea continuă pe care o stârnesc prin gândurile lor şi prin discuţiile fascinante care au creat pretextul, avântul şi materia primă pentru această aventură. Mama mea, Domnica, m-a învăţat cum ştiinţa susţine viaţa şi buna gândire şi m-a convins să cer şi de la tehnologie acelaşi lucru.

Împreună cu Andrei Pitiş şi cu Vladimir Oane am sucit şi răsucit tehnologii, ideologii şi viitoruri posibile pe toate feţele. Discuţiile cu ei mi-au dat impulsul să scriu multe dintre paginile cărţii. Profesorii Nicolae Ţăpuş şi Adina Florea mi-au fost mereu alături, din primul an de facultate, şi au construit reperele intelectuale şi morale spre care m-am orientat. Împreună cu Răzvan Deaconescu şi cu Daniel Rosner am gândit şi răzgândit proiecte, produse, echipe şi comunităţi, într-un studiu aplicat al interacţiunii prieteniei cu tehnologia. Cu Flavia Oprea şi Dan Ciocîrlan am zigzagat toate graniţele dintre mediul universitar şi cel amical şi am băut cantităţi impresionante de cafea. Flavia Husar şi George Sârbu m-au conectat cu noi lumi şi m-au provocat să văd viaţa prin lentile neobişnuite mie. Le mulţumesc Andreei Budeanu, Ceciliei Ciocîrlan şi lui Andrei Stănescu pentru gândurile pe marginea cărţii, lui Vlad Zografi pentru sprijinul, sfaturile şi încurajările pe care mi le-a dat în momente necesare şi lui Alexandru Anghel pentru aducerea cărţii la forma sa cea mai bună.

La final îi mulţumesc vinovatului principal, Constantin Vică. Prin provocările lui şi conversaţii alene, prin confruntările de idei şi experienţe din masteratul de etică aplicată al Universităţii din Bucureşti, m-a convins că putem redefini chiar temeiurile cele mai profunde ale relaţiei noastre cu societatea digitală.

Creditele tabelului și figurilor

Tabelul: Analiză proprie a trăsăturilor Big Data menţionate în postarea „The 42 V’s of Big Data and Data Science“, disponibilă online la https://www.kdnuggets.com/2017/04/42-vs-big-data-data science.html.

Figura 1: YBSLE, imagine publicată pe Wikimedia Commons, disponibilă online la https://commons.wikimedia.org/wiki/File:10_Largest_Corporations_by_Market_Capitalization.png

Figura 2: Statista; grafic disponibil online la https://www.statista.com/chart/21584/gafam-revenue-growth/

Figura 3: Ilustraţie de Gordon Johnson, publicată pe Pixabay, disponibilă online la https://pixabay.com/vectors/code-programming-head-computer-2858768.

Figura 5: Grafic realizat de autor pe baza datelor Statista, disponibile online la https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide.

Figura 6: Grafic realizat de autor pe baza datelor Statista, disponibile online la https://www.statista.com/statistics/277963/ facebooks-quarterly-global-revenue-by-segment.

Figura 7: Statista, grafic disponibil online la https://www.statista. com/chart/18805/highest-penalties-in-privacy-enforcement-actions-worldwide.

Figura 8: Statista, grafic disponibil online la https://www.statista. com/chart/25691/highest-fines-for-gdpr-breaches.

Figura 9: Statista, disponibil online la https://www.statista.com/chart/24748/youtube-advertising-revenue-vs-netflix-revenue.

Figura 10: Grafic realizat de autor pe baza datelor Board of Governors of the Federal Reserve System (US), Student Loans Owned and Securitized [SLOAS], obţinute pe platforma FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/SLOAS, accesat la 5 noiembrie

2021.

Figura 11: Our World in Data, disponibil online la https://ourworldindata.org/us-life-expectancy-low.

Figura 12: Our World in Data, disponibil online la https://ourworldindata.org/grapher/death-rate-from-opioid-use.

Figura 13: Our World in Data, disponibil online la https://ourworldindata.org/grapher/prison-population-rate.

Figura 14: Wikimedia.org, preluat de la November Coalition Sentence; disponibil online la http://november.org/graphs/Americans.gif, Public Domain, https://commons.wikimedia.org/w/index.php? curid= 3854972.

Figura 15: Infograficul Statista, conform datelor Opensecrets.org; disponibil online la https://www.statista.com/chart/22967/overall-election-spending-by-cycle-and-party-affiliation.

Figura 16: Disponibil online la https://theconversation.com/ the-scale- of-us-election-spending-explained-in-five-graphs-130651; cheltuielile au fost împărțite la populațiile statelor din anii respectivi (conform Google).

Figura 17: Our World in Data, disponibil online la https://ourworldindata.org/grapher/inequality-of-incomes-chartbook.

Figura 18: Our World in Data, disponibil online la https://ourworldindata.org/grapher/share-of-top-1-in-pre-tax-national-income.

Figura 19: Our World in Data, disponibil online la https://ourworldindataorg/grapher/gini-index-around-2015-1990-2015-countries-vs-gini-index-around-1990-1990-2015-countries.

Figura 20: Our World in Data, disponibil online la https://ourworldindata.org/the-global-decline-of-extreme-poverty-was-it-only-china.

Figura 21: Our World in Data, disponibil online la https://ourworldindata.org/grapher/share-of-population-in-extreme-poverty.

Figura 22: Infografic Statista, disponibil online la https://lb-aps-frontend.statista.com/chart/23502/market-shares-baidu-alibaba-tencent.

Figura 23: Infografic Statista, disponibil online la https://www. statista. com/chart/22677/the-age-of-the-tech-giants.

Figura 24: Grafic realizat de autor pe baza datelor din United States Census Bureau, disponibile online la https://www.census.gov/data/tables/time-series/ demo/income-poverty/historical-inco-me-house-holds.html.

Figura 26: Captură de ecran publicată în Web Design Museum, https:// www.webdesignmuseum.org/timeline/altavista-1996.

Figura 27: Web Design Museum, Google Timeline, disponibil online la https://www.webdesignmuseum.org/timeline/google-2021) și captură proprie realizată la 13 noiembrie 2021.

Figura 28: Statista, disponibil online la https://www.statista.com/chart/ 23848/most-read-wikipedia-articles-2020.

Figura 29: Statista, disponibil online la https://cdn.statcdn.com/Infographic/images/normal/19058.jpeg.

Figura 30: Statista, disponibil online la https://www.statista.com/ chart/25443/estimated-amount-of-data-created-on-the-internet-in-one-minute.

Figura 31: Our World in Data, disponibil online la https://ourworldindata.org/co2-emissions.

Figura 32: IGBP – International Geosphere Biosphere Programme, disponibil online la http://www.igbp.net/globalchange/greatacceleration.4.1b8ae20512db692f2a680001630.html.

Figura 33: Rezultatele obținute de autor prin căutare pe Google Trends, accesat pe 1 noiembrie 2021.

Figura 34: Rezultatele obținute de autor prin căutare pe Google Trends, accesat pe 1 noiembrie 2021.

Figura 35a: Wikimedia, Rob984; Europe orthographic Caucasus Urals boundary, CC BY-SA 4.0, https://commons. wikimedia.org/ w/index. php?curid=41218742.

Figura 35b: Wikipedia, https://en.wikipedia.org/wiki/File:Global_European_Union.svg.

Figura 35c: Ilustrație de Danlaycock; European Economic Area.svg by Kolja21, Fry1989, CC BY-SA 2.5, https://commons. wikimedia. org/ w/index.php?curid=32132627. Disponibil online în Wikipedia, https://en.wikipedia.org/wiki/European_Economic_Area.

Figura 36: Grafic realizat de autor pe baza datelor UE centralizate de EUobserver.
Figura 37: Grafic realizat de autor pe baza datelor UE centralizate de EUobserver.
Figura 38: Google Trends, accesat pe 10 noiembrie 2021.

Note

INTRODUCERE

1. G. Burke, M. Mendoza, J. Linderman, M. Tarm, „How AI-Powered Tech Landed Man in Jail with Scant Evidence“, Associated Press, publicat online în August 2021, https://apnews.com/article/artificial-intelligence-algorithm-technology-police-crime-7e3345485aa668c97606d4b54f9b6220; Serginho Roosblad, video: „ShotSpotter landed man in jail with scant evidence“, *Associated Press*, publicat online în august 2021: https://apnews.com/article/ap-video-549053630664.
2. M.L. Doucette *et al*., „Impact of ShotSpotter Technology on Firearm Homicides and Arrests Among Large Metropolitan Counties: a Longitudinal Analysis, 1999–2016“, *Journal of Urban Health* (2021), pp. 1–13.
3. Imperva, Bad Bot Report: Bad Bots Strike Back, 2020, https://www.imperva.com/resources/reports/Imperva_BadBot_Report_V2.0.pdf.

1. TEHNOLOGII DIGITALE, CAPITALISM ȘI MODELE SOCIALE: SUA, UE ȘI CHINA

1. A.J. Wood, „Algorithmic Management Consequences for Work Organisation and Working Conditions“, JRC Technical Report, 2021, https://ec.europa.eu/jrc/sites/default/files/jrc124874.pdf.
2. K. Roose, *Futureproof: 9 Rules for Humans in the Age of Automation*, Random House, NY, 2021.
3. M. Waltzer, *Spheres of Justice: A Defense of Pluralism and Equality*, Basic Books, NY, 1983.
4. Contribuitorii Wikipedia, Cambridge Analytica, https://en.wikipedia.org/wiki/Cambridge_Analytica, accesat la 10 aprilie 2021; C. Kang, S. Frenkel, „Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users“, *The New York Times*,

publicat online la 4 aprilie 2018, https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html; R. Rathi, „Effect of Cambridge Analytica's Facebook ads on the 2016 US Presidential Election. *Towards Data Science*", publicat în 2019, https://towardsdatascience.com/ effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d, accesat la 10 aprilie 2021.

5. Statista, Number of monthly active Facebook users worldwide as of 2nd quarter 2021, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/.
6. K. Hao, „How Facebook got addicted to spreading misinformation", *MIT Technology Review*, publicat online în martie 2021, https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation.
7. K. Hao, „How Facebook got addicted to spreading misinformation", *MIT Technology Review*, publicat online în martie 2021, https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation; K. Hao, „The Facebook whistleblower says its algorithms are dangerous. Here's why", *MIT Technology Review*, publicat online în octombrie 2021, https://www.technologyreview.com/2021/10/05/1036519/facebook-whistleblower-frances-haugen-algorithms.
8. Statista, Facebook's global revenue as of 2nd quarter 2021, by segment, publicat în 2021, https://www.statista.com/statistics/277963/facebooks-quarterly-global-revenue-by-segment.
9. K. Hao, „How Facebook got addicted to spreading misinformation", *MIT Technology Review*, publicat online în martie 2021, https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation.
10. K. Hao, „The Facebook whistleblower says its algorithms are dangerous. Here's why", *MIT Technology Review*, publicat online în octombrie 2021, https://www.technologyreview.com/2021/10/05/1036519/ facebook-whistleblower-frances-haugen-algorithms.

11. D. Milmo, „How losing a friend to misinformation drove Facebook whistleblower“, *The Guardian*, publicat online în 2021, https://www.theguardian.com/technology/2021/oct/04/how-friend-lost-to-misinformation-drove-facebook-whistleblower-frances-haugen.
12. M. Friedman, „A Friedman doctrine – The Social Responsibility Of Business Is to Increase Its Profits“, *The New York Times*, 1970, https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html.
13. „Facebook oversight board to review system that exempts elite users“, *The Guardian*, publicat online la 22 septembrie 2021, https://www.theguardian.com/technology/2021/sep/21/facebook-xcheck-system-oversight-board-review.
14. G. Lopez, „Pizzagate, the fake news conspiracy theory that led a gunman to DC’s Comet Ping Pong, explained“, *Vox*, publicat online la 8 decembrie 2016, https://www.vox.com/policy-and-politics/2016/12/5/13842258/pizzagate-comet-ping-pong-fake-news.
15. G. Wells *et. al.*, „Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show“, *The Wall Street Journal*, 14 septembrie 2021, https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739, accesat la 6 noiembrie 2021.
16. G. Wells *et. al.*, „Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show“, *The Wall Street Journal*, 14 septembrie 2021, https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739, accesat la 6 noiembrie 2021.
17. J.C., Wong, „Facebook to be fined $5bn for Cambridge Analytica privacy violations – reports“, *The Guardian*, publicat online în 2019, https://www.theguardian.com/technology/2019/jul/12/facebook-fine-ftc-privacy-violations.
18. Peter Kafka, „Facebook will pay the US government a $5 billion fine for privacy failures – but it won’t have to change the way it does business“, *Vox*, publicat online în 2019. https://www.vox.com/recode/2019/7/24/20708359/facebook-ftc-settlement-criticism-5-billion-privacy-

review-antitrust-mark-zuckerberg.
19. K. Breuninger, L. Feiner, „January 6 committee subpoenas Google, Facebook, Twitter and Reddit in probe of Capitol attack“, *CNBC*, publicat online în 2022, https://www.cnbc.com/2022/01/13/january-6-committee-subpoenas-google-facebook-twitter-and-reddit-in-probe-of-capitol-attack.html.
20. C. Timberg, E. Dwoskin, R. Albergotti, „Inside Facebook, Jan. 6 violence fueled anger, regret over missed warning signs. A trove of internal documents turned over to the SEC provides new details of the social media platform’s role in fomenting the storming of the U.S. Capitol“, *Washington Post*, publicat online în 2021, https://www.washingtonpost.com/technology/2021/10/22/jan-6-capitol-riot-facebook.
21. „The Facebook Files. A *Wall Street Journal* Investigation“, *The Wall Street Journal*, publicat online în 2021, https://www.wsj.com/articles/the-facebook-files-11631713039.
22. J.E. Solsman, „YouTube’s AI is the puppet master over most of what you watch“, *CNET*, publicat în 2018, https://www.cnet.com/news/youtube-ces-2018-neal-mohan.
23. L. Munn, „Angry by design: toxic communication and technical architectures“, *Nature – Humanities and Social Sciences Communications*, 2020, 7 (1), pp. 1–11, https://www.nature.com/articles/s41599-020-00550-7.
24. K. Roose, „The Making of a YouTube Radical“, *The New York Times*, publicat online în 2019. https://www.nytimes.com/interactive/ 2019/ 06/08/technology/youtube-radical.html; P. Martineau, „YouTube Removes More Videos but Still Misses a Lot of Hate“, *Wired*, publicat online în 2019, https://www.wired.com/story/youtube-removes-videos-misses-hate.
25. N. Lomas, „YouTube’s recommender AI still a horror show, finds major crowdsourced study“, *TechCrunch*, publicat în 2021, https://techcrunch.com/2021/07/07/youtubes-recommender-ai-still-a-horrorshow-finds-major-crowdsourced-study; Statista, „Number of YouTube users worldwide from 2016 to 2021“, publicat online în 2021,

https://www.statista.com/statistics/805656/number-youtube-viewers-worldwide.
26. A.J. Hess, „How student debt became a $1.6 trillion crisis“, *CNBC*, publicat online în 2020, https://www.cnbc.com/2020/06/12/how-student-debt-became-a-1point6-trillion-crisis.html.
27. D.M. Johnson, „What Will It Take to Solve the Student Loan Crisis?“, *Harvard Business Review*, publicat online în 2019, https://hbr.org/ 2019/09/what-will-it-take-to-solve-the-student-loan-crisis.
28. A.J. Hess, „How student debt became a $1.6 trillion crisis“, *CNBC*, publicat online în 2020, https://www.cnbc.com/2020/06/12/how-student-debt-became-a-1point6-trillion-crisis.html.
29. D. Oks, L. Herrine, *Why American College Is So Expensive*, 2021, https:// www.youtube.com/watch?v=R58Si78N9i4.
30. A.J. Hess, „The U.S. has a record-breaking $1.73 trillion in student debt – borrowers from these states owe the most on average“, *CNBC*, publicat online în 2021, https://www.cnbc.com/2021/09/09/america-has-1point73-trillion-in-student-debtborrowers-from-these-states-owe-the-most.html.
31. D.M. Johnson, „What Will It Take to Solve the Student Loan Crisis?“, *Harvard Business Review*, publicat online în 2019, https://hbr.org/ 2019/09/what-will-it-take-to-solve-the-student-loan-crisis.
32. E. Zaagman, „The casualties of China’s education crackdown“, *TechCrunch*, publicat online în 2021, https://techcrunch.com/ 2021/09/ 22/the-casualties-of-chinas-education-crackdown.
33. A. Lentacker, „How the Purdue opioid settlement could help the public understand the roots of the drug crisis“, *The Conversation*, publicat online în 2021, https://theconversation.com/how-the-purdue-opioid-settlement-could-help-the-public-understand-the-roots-of-the-drug-crisis-166701.
34. E.A. Shipton, E.E. Shipton, A.J. Shipton, „A Review of the Opioid Epidemic: What Do We Do About It?“, *Pain Therapy*, 2018, 7 (1), pp. 22–36.
35. D. Selman, P. Leighton, *Punishment for Sale: Private Prisons, Big Business, and the Incarceration Binge*, Rowman & Littlefield Publishers, Lanham, 2010.

36. A. Friedmann, „The Societal Impact of the Prison Industrial Complex, or Incarceration for Fun and Profit – Mostly Profit“, *Prison Legal News*, publicat online în 2012, https://www.prisonlegalnews.org/news/2012/jan/15/the-societal-impact-of-the-prison-industrial-complex-or-incarceration-for-fun-and-profitmostly-profit.
37. K. Evers-Hillstrom *et al*., *More Money, Less Transparency: A Decade under Citizens United*, 2020, https://www.opensecrets.org/news/reports/a-decade-under-citizens-united.
38. W.C.R. Horncastle, „The scale of US election spending explained in five graphs“, *The Conversation*, publicat online în 2020, https://theconversation.com/the-scale-of-us-election-spending-explained-in-five-graphs-130651.
39. W.C.R. Horncastle, „The scale of US election spending explained in five graphs“, *The Conversation*, publicat online în 2020, https://theconversation.com/the-scale-of-us-election-spending-explained-in-five-graphs-130651.
40. A. Breeden, „Ex-President Sarkozy Convicted for Campaign Spending Violations“, *The New York Times*, publicat online în 2021, https://www.nytimes.com/2021/09/30/world/europe/france-sarkozy-campaign-spending.html.
41. D. Grathwohl, „The scandal that rocked the government of Helmut Kohl“, *Deutsche Welle*, publicat online în 2004, https://www.dw.com/en/the-scandal-that-rocked-the-government-of-helmut-kohl/a-5137950.
42. T. Piketty, *Capital et idéologie*, Seuil, 2019; versiunea citată este cea în engleză, *Capital and Ideology*, Belknap Press, London, 2020, p. 33.
43. *Ibid*., pp. 40–42.
44. S. Morrison, „California’s new privacy law, explained“, *Vox*, publicat online la 30 decembrie 2019, https://www.vox.com/recode/ 2019/12/30/21030754/ccpa-2020-california-privacy-law-rights-explained.
45. Mark Coeckelbergh, Thomas Metzinger, „Europe needs more guts when it comes to AI ethics“, *Tagesspiel*, 2020, https://background.tagesspiegel.de/digitalisierung/europe-needs-more-guts-when-it-comes-to-ai-ethics.

46. Constantin Vica, Cristina Voinea, Radu Uszkai, „The emperor is naked: Moral diplomacies and the ethics of AI“, *Információs Társadalom* 21, 2 (2021).
47. V Ni, „The party’s over: China clamps down on its tech billionaires“, publicat la 21 august 2021, *The Guardian*, https://www.theguardian.com/business/2021/aug/21/the-partys-over-china-clamps-down-on-its-tech-billionaires.
48. The Economist Editors, „Xi Jinping’s assault on tech will change China’s trajectory“, *The Economist*, publicat online în august 2021, https:// www.economist.com/leaders/2021/08/14/xi-jinpings-assault-on-tech-will-change-chinas-trajectory.
49. E. Zaagman, „The casualties of China’s education crackdown, *TechCrunch*, publicat online în 2021, https://techcrunch.com/ 2021/09/ 22/the-casualties-of-chinas-education-crackdown.
50. L. Clarke, „Why is China cracking down on its tech industry?“, *Tech Monitor*, publicat online în septembrie 2021, https://techmonitor.ai/policy/big-tech/chinese-tech-regulation-alibaba-ant-group-tencent.
51. V Ni, „The party’s over: China clamps down on its tech billionaires“, publicat la 21 august 2021, *The Guardian*, https://www.theguardian.com/business/2021/aug/21/the-partys-over-china-clamps-down-on-its-tech-billionaires; L. Clarke, „Why is China cracking down on its tech industry?“, *Tech Monitor*, publicat online în septembrie 2021, https://techmonitor.ai/policy/big-tech/chinese-tech-regulation-alibaba-ant-group-tencent.
52. A.J. Wood, „Algorithmic Management Consequences for Work Organisation and Working Conditions“, *JRC Technical Report*, 2021, https:// ec.europa.eu/jrc/sites/default/files/jrc124874.pdf; The Economist Editors, „Xi Jinping’s assault on tech will change China’s trajectory“, *The Economist*, publicat online în august 2021, https://www.economist.com/leaders/2021/08/14/xi-jinpings-assault-on-tech-will-change-chinas-trajectory; The Economist Editors, „What tech does China want?“, *The Economist*, publicat online în august 2021, https:// www. economist.com/business/what-tech-does-china-

want/21803410; S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books, NY, 2019.

2. PREAPLINUL CUNOAȘTERII ȘI MINUNATA LUME NOUĂ A PUTERII ALGORITMICE

1. J. Vincent, „The AI oracle of Delphi uses the problems of Reddit to offer dubious moral advice“, *The Verge*, publicat online în octombrie 2021, https://www.theverge.com/2021/10/20/22734215/ai-ask-delphi-moral-ethical-judgement-demo.
2. Redacția *Futurism*, „Scientists Built an AI to Give Ethical Advice, But It Turned Out Super Racist“, *Futurism*, publicat online în octombrie 2021, https://futurism.com/delphi-ai-ethics-racist.
3. Liwei Jiang, „Towards Machine Ethics and Norms“, *Medium*, publicat online în noiembrie 2021, https://medium.com/ai2-blog/towards-machine-ethics-and-norms-d64f2bdde6a3.
4. C. Vică, „Politica zilelor noastre rămâne un mod încă primitiv de a camufla violența“, *Scena 9*, publicat online în 2016, https://www.scena9.ro/article/post-adevar-constantin-vica.
5. C. Voinea, C. Vică, E. Mihailov, J. Savulescu, „The Internet as Cognitive Enhancement“, *Science and Engineering Ethics*, 2020, 26(4), pp. 2345–2362.
6. T. Piketty, *Capital and Ideology*, Belknap Press, Cambridge, 2020.
7. J. Mislinski, „Updated: U.S. Household Incomes – A 50+ Year Perspective“, *Advisor Perspectives*, publicat online în 2020, https://www.advisorperspectives.com/dshort/updates/2020/09/17/updated-u-s-household-incomes-a-50-year-perspective.
8. M. Sainato, „«I’m not a robot»: Amazon workers condemn unsafe, grueling conditions at warehouse“, *The Guardian*, publicat online în 2020, https://www.theguardian.com/technology/2020/feb/05/amazon-workers-protest-unsafe-grueling-conditions-warehouse.
9. V. Dubal, M. Whittaker, „Uber drivers are being forced to choose between risking Covid-19 or starvation“, *The Guardian*, publicat online în 2020, https://www.theguardian.com/technology/2020/mar/25/uber-lyft-gig-economy-coronavirus.

10. C. Newton, „The Trauma Floor“, *The Verge*, publicat online în 2019, https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.
11. A. Stevenson, „Facebook Admits It Was Used to Incite Violence in Myanmar“, *The New York Times*, publicat online în 2018, https://www.nytimes.com/2018/11/06/technology/myanmar-facebook.html.
12. Human Rights Watch, „«Break Their Lineage, Break Their Roots». China’s Crimes against Humanity Targeting Uyghurs and Other Turkic Muslims“, publicat online în 2021, https://www.hrw.org/report/2021/04/19/break-their-lineage-break-their-roots/chinas-crimes-against-humanity-targeting.
13. K. Paul, „«Breeding grounds for radicalization»: Capitol attack panel signals loss of patience with big tech“, *The Guardian*, publicat online în 2022, https://www.theguardian.com/us-news/2022/jan/13/capitol-attack-subpoena-google-facebook-twitter.
14. A.M. Budeanu, D. Rosner, „Big Data as Capital. A Case Study on the Innovation Labs Tech Accelerator“, *23rd International Conference on Control Systems and Computer Science* (CSCS), IEEE, 2021, pp. 469–475.
15. W. Oremus, „Google’s Big Break“, *Slate*, publicat online în 2013, https://slate.com/business/2013/10/googles-big-break-how-bill-gross-goto-com-inspired-the-adwords-business-model.html.
16. W. Oremus, „Google’s Big Break“, *Slate*, publicat online în 2013, https://slate.com/business/2013/10/googles-big-break-how-bill-gross-goto-com-inspired-the-adwords-business-model.html.
17. G. Fox, *Google Business Model – A Hidden Revenue Model*, publicat în 2019, https://www.garyfox.co/google-business-model-explained.
18. Zuboff S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile Books, NY, 2019; S. Zuboff, „The Secrets of Surveillance Capitalism“, *Frankfurter Allgemeine Zeitung*, publicat online în 2016, https://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-secrets-of-surveillance-capitalism-14103616.html.

19. A. Palmer, „What the EU's investigation of Amazon means for U.S. antitrust probes“, *CNBC*, publicat în 2020, https://www.cnbc.com/2020/II/II/eu-investigates-amazon-what-it-means-for-us-antitrust-probes.html; M. Stoller, „At last, the regime that enabled Amazon's monopoly power is crumbling“, *The Guardian*, publicat la 4 februarie 2021, https://www.theguardian.com/commentisfree/2021/feb/04/amazon-jeff-bezos-monopoly-power-antitrust, accesat la 23 octombrie 2021.
20. *European Commission*, Competition Policy, publicat în 2021.
21. J. Nadler, D.N. Cicilline, *Investigation of Competition in Digital Markets. Majority Staff Report and Recommendations*, 2020, https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf?utm_campaign=4493-519.
22. D.A. Smith, „Situating Wikipedia as a health information resource in various contexts: A scoping review“, *PLOS ONE*, publicat online în 2020, https://doi.org/10.1371/journal.pone.0228786.
23. N.J. Temple, „Wikipedia Articles on Nutrition: Are they Accurate and Complete?“, *Current Nutrition & Food Science*, 2020. 16 (2), pp. 237–240.
24. L. Brooks, A. Hern, „Shock an aw: US teenager wrote huge slice of Scots Wikipedia“, *The Guardian*, publicat online în 2020, https://www.theguardian.com/uk-news/2020/aug/26/shock-an-aw-us-teenager-wrote-huge-slice-of-scots-wikipedia.
25. A. Daub, *What Tech Calls Thinking: An Inquiry into the Intellectual Bedrock of Silicon Valley*, Farrar, Straus and Giroux, NY, 2020.
26. Statista, Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2030, publicat în 2021, https://www.statista.com/ statistics/1183457/iot-connected-devices-worldwide/.
27. B. Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-Connected World*, W.W. Norton & Company, NY, 2018.
28. J. Su, V.D. Vargas, S. Kouichi, „One pixel attack for fooling deep neural networks“, *IEEE Transactions on Evolutionary Computation*, 5 (23), publicat online în 2019, https://arxiv.org/abs/1710.08864.

29. D. Goodin, „Researchers trick Tesla Autopilot into steering into oncoming traffic“, *Ars Technica*, publicat online la 4 februarie 2019, https://arstechnica.com/information-technology/2019/04/researchers-trick-tesla-autopilot-into-steering-into-oncoming-traffic, accesat la 23 Octombrie 2021.
30. A. Marshall, „A Fatal Crash Renews Concerns Over Tesla’s «Autopilot» Claim“, *Wired*, publicat online la 20 aprilie 2021, https://www.wired.com/story/fatal-crash-renews-concerns-teslas-autopilot, accesat la 23 octombrie 2021.
31. A. Greenberg, „Split-Second «Phantom» Images Can Fool Tesla’s Autopilot“, *Wired*, publicat online la 10 noiembrie 2020, accesat la 23 octombrie 2021.
32. D. Goodin, „Researchers trick Tesla Autopilot into steering into oncoming traffic“, *Ars Technica*, publicat online la 4 februarie 2019, https://arstechnica.com/information-technology/2019/04/researchers-trick-tesla-autopilot-into-steering-into-oncoming-traffic, accesat la 23 Octombrie 2021; S. Samuel, „It’s disturbingly easy to trick AI into doing something deadly“, *Vox*, publicat online în 2019, https://www.vox.com/future-perfect/2019/4/8/18297410/ai-tesla-self-driving-cars-adversarial-machine-learning.
33. R.L. Carneiro, „The transition from quantity to quality: A neglected causal mechanism in accounting for social evolution“, *Proceedings of the National Academy of Sciences of the United States of America*, 2000, 97 (23), pp. 12926–12931.
34. S.K. Mun *et al.*, „Artificial Intelligence for the Future Radiology Diagnostic Service“, *Frontiers in Mollecural Bioscience*, publicat online în 2021, https://doi.org/10.3389/fmolb.2020.614258; H. Murphy, „Archaeologists vs. Computers: A Study Tests Who’s Best at Sifting the Past“, *The New York Times*, publicat online în 2021.
35. B. Marr, „13 Mind-Blowing Things Artificial Intelligence Can Already Do Today“, *Forbes*, Publicat online în 2019, https://www.forbes.com/sites/bernardmarr/2019/11/11/13-mind-blowing-things-artificial-intelligence-can-already-do-today.
36. M. Constantinescu, C. Voinea, R. Uszkai, C. Vică, „Understanding

responsibility in Responsible AI. Dianoetic virtues and the hard problem of context“, *Ethics and Information Technology*, publicat online la 30 septembrie 2021, pp. 1–12.
37. S. Vallor, „Moral Deskilling and Upskilling in a New Machine Age: Reflections on the Ambiguous Future of Character“, *Philosophy of Technology*, 2015, 28, pp. 107–124.
38. The Economist Editors, „Posen A, Jackson T. Slower Growth – Disaster or Blessing?“, *The Economist: The World If*, publicat online în 2015, https://worldif.economist.com/article/12121/debate.
39. T. Jackson, P.A. Victor, „Unraveling the claims for (and against) green growth“, *Science*, 2019, 366, pp. 64–68.
40. C. Rosol, *et al.*, „On the age of computation in the epoch of humankind“, *Nature Outlook*, publicat în 2018, https://www.nature.com/articles/d42473-018-00286-8, accesat la 16 octombrie 2021.
41. W. Knight, „AI’s Smarts Now Come With a Big Price Tag“, *Wired*, publicat online în 2021, https://www.wired.com/story/ai-smarts-big-price-tag; J. Snow, „How artificial intelligence can tackle climate change“, *National Geographic*, publicat online în 2019 https://www.nationalgeographic.com/environment/article/artificial-intelligence-climate-change; K. Hao, „Here are 10 ways AI could help fight climate change“, *MIT Technology Review*, publicat online în 2019, https://www. technologyreview.com/2019/06/20/134864/ai-climate-change-machine-learning.
42. G. Conole, M. Dyke, „What are the affordances of information and communication technologies?“, *ALT-j*, 2004, 12 (2), pp. 113–124.

3. EXCESUL CUNOAȘTERII ÎN SOCIETATEA ALGORITMICĂ

1. M. Allen, „You Snooze, You Lose: Insurers Make The Old Adage Literally True“, *ProPublica*, publicat online la 21 noiembrie 2018, https:// www.propublica.org/article/you-snooze-you-lose-insurers-make-the-old-adage-literally-true.
2. A. Alexander, „Is Your Sleep Apnea Machine Snitching to Your Insurer?“, *Bloomberg Law*, publicat online la 20 martie 2019, https:// news.bloomberglaw.com/health-law-and-business/is-your-sleep-apnea-

machine-snitching-to-your-insurer.
3. D. Kravitz, M. Allen, „Your Medical Devices Are Not Keeping Your Health Data to Themselves“, *ProPublica*, publicat online la 21 noiembrie 2018, https://www.propublica.org/article/your-medical-devices-are-not-keeping-your-health-data-to-themselves.
4. M.S. Rusu, *Delicii prozaice: Bârfa ca practică socială a vieții cotidiene*, Institutul European, Iași, 2020, capitolul 8.
5. O. Schwartz, „Digital Ads are Starting to Feel Psychic“, *The Outline*, publicat online la 13 iulie 2018, https://theoutline.com/post/5380/targeted-ad-creepy-surveillance-facebook-instagram-google-listening-not-alone, accesat la 23 octombrie 2021.
6. R. Jennings, „Why targeted ads are the most brutal owns“, *Vox*, publicat online la 25 Septembrie 2018, https://www.vox.com/the-goods/2018/9/25/17887796/facebook-ad-targeted-algorithm, accesat la 23 octombrie 2021.
7. Q. Wong, „Facebook’s ad targeting has created a creepy image problem it can’t shake“, *CNET*, publicat la 16 mai 2019, https://www.cnet. com/tech/computing/facebooks-ad-targeting-has-created-a-creepy-image-problem-it-cant-shake, accesat la 23 octombrie 2021.
8. Q. Wong, „Facebook’s ad targeting has created a creepy image problem it can’t shake“, *CNET*, publicat la 16 mai 2019, https://www.cnet. com/tech/computing/facebooks-ad-targeting-has-created-a-creepy-image-problem-it-cant-shake, accesat la 23 octombrie 2021.
9. T. Brain, S. Lavigne, „The New Organs“, publicat în 2018, https://neworgans.net/about.html, accesat la 23 octombrie 2021.
10. A.G. Martinez, „Facebook’s Not Listening Through Your Phone. It Doesn’t Have To“, *Wired*, publicat online la 11 octombrie 2017, https://www.wired.com/story/facebooks-listening-smartphone-microphone, accesat la 30 octombrie 2021; K. Tiffany, „The perennial debate about whether your phone is secretly listening to you, explained“, *Vox*, publicat online la 28 decembrie 2018, https://www.vox. com/the-goods/2018/12/28/18158968/facebook-microphone-tapping-recording-instagram-ads, accesat la 30 octombrie 2021.

11. J. Lucker, S.K. Hogan, T. Bischoff, „Predictably inaccurate. The prevalence and perils of bad big data“, *Deloitte Review*, 2017 (21), https://www2.deloitte.com/content/dam/insights/us/articles/3924_Predictably-inaccurate/DUP_Predictably-inaccurate-reprint.pdf; N. Neumann, C.E. Tucker, T. Whitfield, „Frontiers: How effective is third-party consumer profiling? Evidence from field studies“, *Marketing Science*, 2019, 38 (6), pp. 918–926.
12. A. Fou, „Ad «Relevancy» Is Fiction, And It’s Creepy“, *Forbes*, publicat online la 19 aprilie 2021, https://www.forbes.com/sites/augustinefou/2021/04/19/ad-relevancy-is-fiction-and-its-creepy, accesat la 23 octombrie 2021.
13. J. Marciano, „The global growth of ad fraud in 15 countries, *CHEQ*, publicat în 2021, https://www.cheq.ai/the-global-growth-of-ad-fraud-in-15-countries, accesat la 23 octombrie 2021.
14. A. Storozuk *et al.*, „Got Bots? Practical Recommendations to Protect Online Survey Data from Bot Attacks“, *The Quantitative Methods for Psychology*, 2020, 16 (5), pp. 472–481; J.E.F. Teitcher *et al.*, „Detecting, Preventing, and Responding to «Fraudsters» in Internet Research: Ethics and Tradeoffs“, *The Journal of law, medicine & ethics: a journal of the American Society of Law, Medicine & Ethics*, 2015, 43 (1), 116.
15. A. Fou, „Bots Do Whatever You Pay For, At Scale“, *Forbes*, publicat în 2021, https://www.forbes.com/sites/augustinefou/2021/04/20/bots-do-whatever-you-pay-for, accesat la 23 octombrie 2021.
16. D. Rosner, *Remote Monitoring and Assisted Living Technologies for Ubiquitous Health Care*, Printech, 2021.
17. S. Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books, NY, 2019; S. Zuboff, „The Secrets of Surveillance Capitalism“, *Frankfurter Allgemeine Zeitung*, publicat online în 2016, https://www.faz.net/aktuell/feuilleton/debatten/the-digital-debate/shoshana-zuboff-secrets-of-surveillance-capitalism-14103616.html; J. Naughton, „Interview with Shosana Zuboff: «The goal is to automate us»: welcome to the age of surveillance capitalism“, *The*

Observer, publicat în 2019, https://www.theguardian.com/technology/2019/jan/20/shoshana-zuboff-age-of-surveillance-capitalism-google-facebook.
18. Google, Micro-Moments. Think With Google, publicat în 2021, https://www.thinkwithgoogle.com/marketing-strategies/micro-moments.
19. A. Nadler, M. Crain, J. Donovan, „Weaponizing the Digital Influence Machine: The Political Perils of Online Ad Tech“, *Data* & *Society*, publicat online în 2018, https://datasociety.net/wp-content/ uploads/ 2018/10/DS_Digital_Influence_Machine.pdf.
20. N. Lomas, „Facebook is weaponizing security to erode privacy“, *TechCrunch*, publicat online la 29 septembrie 2018, https://techcrunch. com/2018/09/29/facebook-is-weaponizing-security-to-erode-privacy, accesat la 23 octombrie 2021.
21. N. Lomas, „Facebook is weaponizing security to erode privacy“, *TechCrunch*, publicat online la 29 septembrie 2018, https://techcrunch. com/2018/09/29/facebook-is-weaponizing-security-to-erode-privacy, accesat la 23 octombrie 2021.
22. A. Zwitter, „Big Data ethics“, *Big Data* & *Society*, 2014, 1 (2).
23. C. Duhigg, „How Companies Learn Your Secrets“, *The New York Times*, publicat la 16 februarie 2012, https://www.nytimes.com/2012/02/19/ magazine/shopping-habits.html.
24. M. Kosinski, D. Stillwell, T. Graepel, „Private traits and attributes are predictable from digital records of human behavior“, *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110 (15), 5802-5805.
25. C. Kang, S. Frenkel, „Facebook Says Cambridge Analytica Harvested Data of Up to 87 Million Users“, *The New York Times*, publicat online la 4 aprilie 2018, https://www.nytimes.com/2018/04/04/ technology/ mark-zuckerberg-testify-congress.html.
26. R. Rathi, „Effect of Cambridge Analytica’s Facebook ads on the 2016 US Presidential Election“, *Towards Data Science*, publicat în 2019, https://towardsdatascience.com/effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d, accesat la 10 aprilie 2021.

27. T. McNeil, „Did Cambridge Analytica Sway the Election? Tufts political scientist Eitan Hersh offers a different take on the controversial firm at a Senate Judiciary Committee hearing“, *TuftsNow*, publicat la 17 mai 2018, https://now.tufts.edu/articles/did-cambridge-analytica-sway-election.
28. D. Sumpter, *Outnumbered*, Bloomsbury, London, 2018.
29. R. Rathi, „Effect of Cambridge Analytica’s Facebook ads on the 2016 US Presidential Election“, *Towards Data Science*, publicat în 2019, https://towardsdatascience.com/effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d, accesat la 10 aprilie 2021.
30. A. Zwitter, „Big Data ethics“, *Big Data & Society*. 2014, 1 (2).
31. M. Barbaro, T. Zeller, „A Face Is Exposed for AOL Searcher No. 4417749“, *The New York Times*, publicat la 9 august 2006, https://www.nytimes.com/2006/08/09/technology/09aol.html.
32. D. McCullagh, „AOL’s disturbing glimpse into users’ lives“, *CNET*, publicat în 2006, https://www.cnet.com/news/aols-disturbing-glimpse-into-users-lives, accesat la 16 aprilie 2021.
33. K. Wagner, „This is how Facebook collects data on you even if you don’t have an account“, *Vox*, publicat online la 20 aprilie 2018, https://www.vox.com/2018/4/20/17254312/facebook-shadow-profiles-data-collection-non-users-mark-zuckerberg, accesat la 24 octombrie 2021.
34. J. Leibold, „Surveillance in China’s Xinjiang Region: Ethnic Sorting, Coercion, and Inducement“, *Journal of Contemporary China*, 2020, 29 (121), pp. 46–60.
35. J. Bhuiyan, „«There’s cameras everywhere»: testimonies detail far-reaching surveillance of Uyghurs in China“, *The Guardian*, publicat online la 30 septembrie 2021.
36. W. Knight, „Clearview AI Has New Tools to Identify You in Photos“, *Wired*, publicat online la 10 aprilie 2021, https://www.wired.com/story/clearview-ai-new-tools-identify-you-photos, accesat la 24 octombrie 2021.
37. R. Metz, „Think your mask makes you invisible to facial recognition?

Not so fast, AI companies say“, *CNN Business*, publicat online la 12 august 2020, https://edition.cnn.com/2020/08/12/tech/face-recognition-masks/index.html, accesat la 24 octombrie 2021.
38. M. Heikkila, „POLITICO AI: Decoded – What EU capitals think of the AI Act. – Facebook’s content moderation headache“, *Politico*, publicat online la 20 octombrie 2021, https://www.politico.eu/newsletter/ai-decoded/ai-goes-to-school-what-eu-capitals-think-of-the-ai-act-facebooks-content-moderation-headache-2/, accesat la 24 octombrie 2021.
39. C. Fraser, „Target didn’t figure out a teenager was pregnant before her father did, and that one article that said they did was silly and bad“, *Medium*, publicat în 2020, https://medium.com/@colin. fraser/ target-didnt-figure-out-a-teen-girl-was-pregnant-before-her-father-did-a6be13b973a5, accesat la 10 aprilie 2020.
40. R. Rathi, „Effect of Cambridge Analytica’s Facebook ads on the 2016 US Presidential Election“, *Towards Data Science*, publicat în 2019, https://towardsdatascience.com/effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d, accesat la 10 aprilie 2021.
41. C. O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, New York, 2017.
42. R. Rathi, „Effect of Cambridge Analytica’s Facebook ads on the 2016 US Presidential Election“, *Towards Data Science*, publicat în 2019, https://towardsdatascience.com/effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d, accesat la 10 aprilie 2021.
43. Knight W., „Clearview AI Has New Tools to Identify You in Photos“, *Wired*, publicat online la 10 aprilie 2021, https://www.wired. com/ story/clearview-ai-new-tools-identify-you-photos, accesat la 24 octombrie 2021.
44. J. Ball, „Costeja González and a memorable fight for the «right to be forgotten»“, *The Guardian*, publicat online la 14 mai 2014, https:// www.theguardian.com/world/blog/2014/may/14/mario-costeja-gonzalez-fight-right-forgotten, accesat la 7 noiembrie 2021.

45. M. Peguera, „No More Right to Be Forgotten for Mr. Costeja, Says Spanish Data Protection Authority“, The Center for Internet and Society, Stanford University, publicat la 3 octombrie 2015, http://cyberlaw.stanford.edu/blog/2015/10/no-more-right-be-forgotten-mr-costeja-says-spanish-data-protection-authority, accesat la 7 noiembrie 2021.
46. C. Goldin, C. Rouse, „Orchestrating Impartiality: The Impact of «Blind» Auditions on Female Musicians“, *American Economic Review*, 2000, 90 (4), pp. 715–741; C. Rice, „How blind auditions help orchestras to eliminate gender bias“, *The Guardian*, publicat online la 14 octombrie 2013, https://www.theguardian.com/women-in-leadership/2013/oct/14/blind-auditions-orchestras-gender-bias, accesat la 30 octombrie 2021.
47. A. Yanchur *et al.*, „Computer says No: How the EU's AI laws cause new injustice“, *EUobserver*, publicat online la 23 august 2021, https://euobserver.com/investigations/152695, accesat la 31 octombrie 2021.
48. S. Bozstas, „The childcare benefits scandal: voices of the victims“, *Dutch News*, publicat online în 2021, https://www.dutchnews.nl/news/2021/01/the-childcare-benefits-scandal-voices-of-the-victims, accesat la 31 octombrie 2021.
49. „Methods used by Dutch Tax and Customs Administration unlawful and discriminatory“, Dutch Data Protection Authority, publicat online în 2020,. https://autoriteitpersoonsgegevens.nl/en/news/methods-used-dutch-tax-and-customs-administration-unlawful-and-discriminatory, accesat la 31 octombrie 2021.
50. C. Goldin, C. Rouse, „Orchestrating impartiality: The impact of «blind» auditions on female musicians“, *American Economic Review*, 2000, 90 (4), pp. 715–741.
51. M. Flyverbom *et al.*, „The Governance of Digital Technology, Big Data, and the Internet: New Roles and Responsibilities for Business“, *Business & Society*, 2019, 58 (1), pp. 3–19.
52. P.T. Kim, „Big Data and Artificial Intelligence: New Challenges for Workplace Equality“, *University of Louisville Law Review*, 2018, 57, https://heinonline.org/HOL/Page?

handle=hein.journals/branlaj57&id=325&div=&collection, accesat la 19 aprilie 2021.
53. P.M. Asaro, „AI ethics in predictive policing: From models of threat to an ethics of care“, *IEEE Technology and Society Magazine*, 2019, 38 (2), pp. 40–53.
54. P.M. Asaro, „AI ethics in predictive policing: From models of threat to an ethics of care“, *IEEE Technology and Society Magazine*, 2019, 38 (2), pp. 40–53.
55. Editors, „The Staggering Cost of Training SOTA AI Models“, *Synced Review*, publicat în 2019, https://syncedreview.com/2019/06/27/the-staggering-cost-of-training-sota-ai-models, accesat la 15 aprilie 2021.
56. The Economist Editors, „The cost of training machines is becoming a problem“, *The Economist*, publicat online în 2020, https://www.economist.com/technology-quarterly/2020/06/11/the-cost-of-training-machines-is-becoming-a-problem.
57. K. Hao, „Training a single AI model can emit as much carbon as five cars in their lifetimes“, *MIT Technology Review*, publicat online în 2019, https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes.
58. K. Hao, „Training a single AI model can emit as much carbon as five cars in their lifetimes“, *MIT Technology Review*, publicat online în 2019, https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes.
59. E. Masanet *et al.*, „Recalibrating global data center energy-use estimates“, *Science*, 2020, 367 (6481).
60. The Economist Editors, „Posen A, Jackson T. Slower Growth – Disaster or Blessing?“, *The Economist: The World If*, publicat online în 2015, https://worldif.economist.com/article/12121/debate.
61. Jackson T, Victor PA. Unraveling the claims for (and against) green growth. *Science*. 2019; 366(6468).
62. A. Chen, „The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed“, *Wired*, publicat online în octombrie 2014, https://www.wired.com/2014/10/content-moderation.
63. C. Newton, „The Trauma Floor“, *The Verge*, publicat în online în

februarie 2019. https://www.theverge.com/2019/2/25/18229714/cognizant-facebook-content-moderator-interviews-trauma-working-conditions-arizona.
64. C. Newton, „How Grindr became a national security issue“, *The Verge*, publicat online la 28 martie 2019, https://www.theverge.com/interface/2019/3/28/18285274/grindr-national-security-cfius-china-kunlun-military, accesat la 7 noiembrie 2021.
65. Z. Whittaker, „Grindr sold by Chinese owner after US raised national security concerns“, *TechCrunch*, publicat online la 20 martie 2020, https://techcrunch.com/2020/03/06/grindr-sold-china-national-security, accesat la 7 noiembrie 2021.
66. „Pillar Investigates: USCCB gen sec Burrill resigns after sexual misconduct allegations“, *The Pillar*, publicat online la 20 iulie 2021, https://www.pillarcatholic.com/p/pillar-investigates-usccb-gen-sec, accesat la 31 octombrie 2021.
67. „Location-based apps pose security risk for Holy See“, *The Pillar*, publicat online la 27 iulie 2021, https://www.pillarcatholic.com/p/location-based-apps-pose-security.
68. „Newark archdiocese to investigate app use allegations“, *The Pillar*, publicat online la 23 iulie 2021, https://www.pillarcatholic.com/p/newark-archdiocese-to-investigate.
69. „Grindr fine cut to $7 million in Norway data privacy case“, Reuters, publicat online la 16 decembrie 2021, https://www.euronews.com/next/2021/12/15/grindr-dataprotection-norway
70. J. Valentino-DeVries, „Your Apps Know Where You Were Last Night, and They’re Not Keeping It Secret“, *New York Times*, publicat online la 10 decembrie 2018, https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html.
71. J. Cox, „How the U.S. Military Buys Location Data from Ordinary Apps“, *Motherboard*, publicat online la 16 noiembrie, 2020, https://www.vice.com/en/article/jgqm5x/us-military-location-data-xmode-locate-x.
72. B. Tau, M. Hackman, „Federal Agencies Use Cellphone Location Data for Immigration Enforcement“, *The Wall Street Journal*, https:// www.

wsj.com/articles/federal-agencies-use-cellphone-location-data-for-immigration-enforcement-11581078600.
73. M. Guariglia, „Law Enforcement Purchasing Commercially-Available Geolocation Data is Unconstitutional“, *Electronic Frontier Foundation*, publicat online la 3 decembrie 2020, https://www. eff.org/ deeplinks/ 2020/12/law-enforcement-purchasing-commercially-available-geolocation-data.
74. H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University Press, Stanford, 2010.
75. M. Andrejevic, „The Big Data Divide“, *International Journal of Communication*, 2014, 8 (1), 17.
76. A. Acquisti *et al.*, „Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online“, *ACM Computing Surveys*, 2017, 50 (3), pp. 1–44.
77. A. Utrata, Postare Twitter publicată online la 23 ianuarie 2022, https://twitter.com/AlinaUtrata/status/1485194962027388929?ref_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Etweet.
78. F. Kaltheuner, „I asked an online tracking company for all of my data and here's what I found“, *PI* (Privacy International), publicat în 2018, https://privacyinternational.org/long-read/2433/i-asked-online-tracking-company-all-my-data-and-heres-what-i-found, accesat la 10 aprilie 2021.
79. J.J. Roberts, „Exclusive: For $3, a «robot lawyer» will sue data brokers that don't delete your personal and location info“, *Fortune*, publicat online la 5 martie 2020, https://fortune.com/2020/03/05/delete-location-data-privacy-personal-information-donotpay.

4. DEFICITUL CUNOAȘTERII ÎN SOCIETATEA ALGORITMICĂ

1. T. Simonite, „Facebook Is Everywhere; Its Moderation Is Nowhere Close“, *Wired*, publicat online la 25 octombrie 2021, https://www.wired.com/story/facebooks-global-reach-exceeds-linguistic-grasp, accesat la 7 noiembrie 2021.
2. „The Facebook Files. A *Wall Street Journal* Investigation“, *The Wall Street Journal*, publicat online în 2021, https://www.wsj. com/articles/ the-facebook-files-11631713039.
3. T. Simonite, „Facebook Is Everywhere; Its Moderation Is Nowhere

Close“, *Wired*, publicat online la 25 octombrie 2021, https://www.wired.com/story/facebooks-global-reach-exceeds-linguistic-grasp, accesat la 7 noiembrie 2021.
4. Y. Wijeratne, „Facebook, language and the difficulty of moderating hate speech“, *London School of Economics* blog, publicat la 23 iulie 2020, accesat la 7 noiembrie 2021; M. Fick, P. Dave, „Facebook’s flood of languages leave it struggling to monitor content“, *Reuters*, publicat la 23 aprilie 2019, https://www.reuters.com/article/us-facebook-languages-insight-idUSKCN1RZ0DW.https://blogs.lse.ac.uk/ medialse/ 2020/07/23/facebook-language-and-the-difficulty-of-moderating-hate-speech, accesat la 7 noiembrie 2021.
5. M. Fick, P. Dave, „Facebook’s flood of languages leave it struggling to monitor content“, *Reuters*, publicat la 23 aprilie 2019, https://www.reuters.com/article/us-facebook-languages-insight-idUSKCN1 RZ0DW, accesat la 7 noiembrie 2021.
6. T. Schafer, „The 42 V’s of Big Data and Data Science“, *KDnuggets*, publicat în 2017, https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html, accesat la 10 aprilie 2021.
7. T. Schafer, „The 42 V’s of Big Data and Data Science“, *KDnuggets*, publicat în 2017, https://www.kdnuggets.com/2017/04/42-vs-big-data-data-science.html, accesat la 10 aprilie 2021.
8. M. C. Elish, D. Boyd, „Situating methods in the magic of Big Data and AI“, *Communication monographs*, 2018, 85 (1), pp. 57–80.
9. Danah Boyd, Kate Crawford, „«Six provocations for big data». A decade in internet time: Symposium on the dynamics of the internet and society“, 2011, https://papers.ssrn.com/sol3/ papers.cfm? abstract_id=1926431.
10. A. Zoldan, „More Data, More Problems: Is Big Data Always Right?“, *Wired*, publicat online în 2013, https://www.wired.com/insights/ 2013/05/more-data-more-problems-is-big-data-always-right.
11. P.T. Kim, „Data-Driven Discrimination at Work“, *William & Mary Law Review*, 2017, 58 (3).
12. A. Zoldan, „More Data, More Problems: Is Big Data Always Right?“, *Wired*, publicat online în 2013, https://www.wired.com/insights/

2013/05/more-data-more-problems-is-big-data-always-right.
13. F. Kaltheuner, „I asked an online tracking company for all of my data and here's what I found", *PI* (Privacy International), publicat online în 2018, https://privacyinternational.org/long-read/2433/i-asked-online-tracking-company-all-my-data-and-heres-what-i-found, accesat la 10 aprilie 2021.
14 J. Lucker, S.K. Hogan, T. Bischoff, „Predictably inaccurate. The prevalence and perils of bad big data", *Deloitte Review*, 2017, (21), https://
www2.deloitte.com/content/dam/insights/us/articles/3924_Predictably-inaccurate/DUP_Predictably-inaccurate-reprint.pdf.
15. J. Lucker, S.K. Hogan, T. Bischoff, „Predictably inaccurate. The prevalence and perils of bad big data", *Deloitte Review*, 2017, (21), https://
www2.deloitte.com/content/dam/insights/us/articles/3924_Predictably-inaccurate/DUP_Predictably-inaccurate-reprint.pdf.
16. C.G. Northcutt, A. Athalye, J. Mueller, „Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks", *35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks*, publicat online la 26 martie 2021, http://arxiv.org/ abs/2103.14749, accesat la 12 aprilie 2021.
17. W. Knight, „The Foundations of AI Are Riddled With Errors", *Wired*, publicat online în 2021, https://www.wired.com/story/foundations-ai-riddled-errors.
18. European Parliament and Council of the European Union, The General Data Protection Regulation – Regulation (EU) 2016/679, publicat în 2016, https://gdpr-info.eu.
19. K. Hill, „Wrongfully accused by an algorithm", *The New York Times*, publicat online în 2020, https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html.
20. J. Buolamwini, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, Vol 81, PMLR, 2018, http://proceedings. mlr.press/v81/buolamwini18a.html, accesat la 12

aprilie 2021.
21. J. Snow, „Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots", *American Civil Liberties Union*, publicat în 2018, https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28, accesat la10 aprilie 2021.
22. K. Hill, „Another Arrest, and Jail Time, Due to a Bad Facial Recognition Match", actualizat la 6 ianuarie 2021, *The New York Times*, publicat la 6 ianuarie 2021.
23. C. Garvie, „The Untold Number of People Implicated in Crimes They Didn't Commit Because of Face Recognition", publicat online în 2020, https://www.aclu.org/news/privacy-technology/the-untold-number-of-people-implicated-in-crimes-they-didnt-commit-because-of-face-recognition.
24. R. van Noorden, „The ethical questions that haunt facial-recognition research", *Nature*, 2020, 587, pp. 354–358, https://media.nature. com/original/magazine-assets/d41586-020-03187-3/d41586-020-03187-3.pdf.
25. E. Kwong, M.K. Sofia, B. Hanson, „Why Tech Companies Are Limiting Police Use of Facial Recognition", *National Public Radio*, publicat online în 2021, https://www.npr.org/2021/02/17/968710172/why-tech-companies-are-limiting-police-use-of-facial-recognition.
26. ACM U.S. Technology Policy Committee, „Statement on Principles and Prerequisites for the Development, Evaluation and Use of Unbiased Facial Recognition Technologies", publicat în 2020, https://www.acm.org/binaries/content/assets/public-policy/ustpc-facial-recognition-tech-statement.pdf, accesat la 10 aprilie 2021.
27. H. Ekbia, *et al.*, „Big data, bigger dilemmas: A critical review", *Journal of the Association for Information Science and Technology*, 2015, 66 (8), pp. 1523–1545.
28. S. Schwarzkopf, „Sacred Excess: Organizational Ignorance in an Age of Toxic Data", *Organization Studies*, 2020, 41 (2), pp. 197–217.
29. C. O'Neil, „The Era of Blind Faith in Big Data Must End", *TED Talks*, publicat în 2017, accesat la 10 aprilie 2011.
30. A. Zoldan, „More Data, More Problems: Is Big Data Always Right?",

Wired, publicat în 2013, https://www.wired.com/insights/2013/ 05/more-data-more-problems-is-big-data-always-right.
31. C. O'Neil, „The Era of Blind Faith in Big Data Must End", *TED Talks*, publicat în 2017, accesat la 10 aprilie 2011.
32. M. Janssen, G. Kuk, „The challenges and limits of big data algorithms in technocratic governance", *Government Information Quarterly*, 2016, 33 (3), pp. 371–377.
33. W. Knight, „The Apple Card Didn't «See» Gender – and That's the Problem", *Wired*, publicat online în noiembrie 2019, https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem.
34. W. Knight, „The Apple Card Didn't «See» Gender – and That's the Problem", *Wired*, publicat online în noiembrie 2019, https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem.
35. I.C. Campbell, „The Apple Card doesn't actually discriminate against women, investigators say", *The Verge*, publicat online în martie 2021, https://www.theverge.com/2021/3/23/22347127/goldman-sachs-apple-card-no-gender-discrimination.
36. J. Noel, „On the Varieties of Phronesis", *Educational Philosophy and Theory*, 1999, 31(3), pp. 273–289.
37. J. Noel, „On the Varieties of Phronesis", *Educational Philosophy and Theory*, 1999, 31(3), pp. 273–289.
38. A. Connock, A. Stephen, „We invited an AI to debate its own ethics in the Oxford Union – what it said was startling", *The Conversation*, publicat online la 10 decembrie, 2021, https://theconversation.com/ we-invited-an-ai-to-debate-its-own-ethics-in-the-oxford-union-what-it-said-was-startling-173607.
39. J. Malik, intervenție video citată în W. Knight, „A Stanford Proposal Over AI's «Foundations» Ignites Debate", *Wired*, publicat online la 14 septembrie 2021, https://www.wired.com/story/stanford-proposal-ai-foundations-ignites-debate.
40. E. Ratti, „Phronesis and Automated Science: The Case of Machine Learning and Biology", *Springer*, 2020, pp. 157–172.
41. T. Simonite, „When It Comes to Health Care, AI Has a Long Way to Go", *Wired*, publicat online la 16 ianuarie 2022, https://www.wired.

com/story/health-care-ai-long-way-to-go.
42. D.M. Christensen *et al.*, „Medical Algorithms Are Failing Communities Of Color“, *Health Affairs*, publicat online la 9 septembrie 2021, https://www.healthaffairs.org/do/10.1377/forefront.20210903.976632/full .
43. M. Guariglia, „Police Use of Artificial Intelligence: 2021 in Review“, *Electronic Frontier Foundation*, publicat online la 1 ianuarie 2022, https:// www.eff.org/deeplinks/2021/12/police-use-artificial-intelligence-2021-review.
44. W.D. Smart, „The robot revolution has been rescheduled (until we can debug the sensors): Technical aspects of robotics“, *WeRobot pre-conference Workshop*, 2016, Miami, FL, citat în M.C. Elish, D. Boyd, „Situating methods in the magic of Big Data and AI“, *Communication Monographs*, 2018, 85 (1), pp. 57–80.
45. D. McQuillan, „Data science as machinic neoplatonism“, *Philosophy & Technology*, 2018, 31 (2), pp. 253–272.
46. B. Berendt *et al.*, „Is it Research or is it Spying? Thinking-Through Ethics in Big Data AI and Other Knowledge Sciences“, *KI - Kunstliche Intelligenz*, Springer, 2015, 29 (2), pp. 223–232.
47. A. Smart *et al*., „Why Reliabilism Is Not Enough: Epistemic and Moral Justification in Machine Learning“, *AIES 2020 – Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, publicat online în 2020.
48. C. Voinea, C. Vică, E. Mihailov, J. Savulescu, „The Internet as Cognitive Enhancement“, *Science and Engineering Ethics*, publicat online la 6 aprilie 2020, pp. 1–18.
49. C. Voinea, C. Vică, E. Mihailov, J. Savulescu, „The Internet as Cognitive Enhancement“, *Science and Engineering Ethics*, publicat online la 6 aprilie 2020, pp. 1–18.
50. C. Voinea, C. Vică, E. Mihailov, J. Savulescu, „The Internet as Cognitive Enhancement“, *Science and Engineering Ethics*, publicat online la 6 aprilie 2020, pp. 1–18.
51. M. Frické, „Big data and its epistemology“, *Journal of the Association for Information Science and Technology*, 2015, 66 (4), pp. 651–661.

52. J. Metcalf, K. Crawford, „Where are human subjects in Big Data research? The emerging ethics divide“, *Big Data & Society*, 2016, 3 (1).
53. S. Dominus, „The Covid Drug Wars That Pitted Doctor vs. Doctor“, *The New York Times*, publicat la 5 august 2020, https://www.nytimes. com/2020/08/05/magazine/covid-drug-wars-doctors.html.
54. J. Metcalf, K. Crawford, „Where are human subjects in Big Data research? The emerging ethics divide“, *Big Data & Society*, 2016, 3 (1).
55. J. Metcalf, K. Crawford, „Where are human subjects in Big Data research? The emerging ethics divide“, *Big Data & Society*, 2016, 3 (1).
56. M. v. Hauge *et al.*, „Tagging Banksy: using geographic profiling to investigate a modern art mystery“, *Journal of Spatial Science*, 2016, 61 (1), pp. 185–190.
57. H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*, Stanford University Press, Stanford, 2010.
58. M.C. Benigni *et al.*, „Online extremism and the communities that sustain it: Detecting the ISIS supporting community on Twitter“, *PLOS ONE*, 2017, 12 (12).
59. E. Buchanan, „Considering the ethics of big data research: A case of Twitter and ISIS/ISIL“, Gómez S, ed., *PLOS ONE*, 2017, 12 (12).
60. E. Buchanan, „Considering the ethics of big data research: A case of Twitter and ISIS/ISIL“, Gómez S, ed., *PLOS ONE*, 2017, 12 (12).
61. C. Wang *et al.*, „Facial feature discovery for ethnicity recognition“, *WIREs Data Mining and Knowledge Discovery*, 2019, 9 (1).
62. R. van Noorden, „The ethical questions that haunt facial-recognition research“, *Nature*, 2020, 587, pp. 354–358, https://media.nature. com/ original/magazine-assets/d41586-020-03187-3/d41586-020-03187-3.pdf.
63. C. Voinea, C. Vică, E. Mihailov, J. Savulescu, „The Internet as Cognitive Enhancement“, *Science and Engineering Ethics*, 2020, 26 (4), pp. 2345–2362; C. Voinea, C. Vică, E. Mihailov, J. Savulescu, „The Internet as Cognitive Enhancement“, *Science and Engineering Ethics*, publicat online la 6 aprilie 2020, pp. 1–18.
64. C. Voinea, „We are stupid and contagious: infodemia și falsele ei remedii“, *Scena 9*, publicat online în 2020, https://www.scena9.ro/article/infodemie-dezinformare-coronavirus.

65. L. Alexander, C. Silverman, „How Teens In The Balkans Are Duping Trump Supporters With Fake News“, *BuzzFeed News*, publicat online la 4 noiembrie 2016, https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo.
66. D. Kurtzleben, „With «Fake News», Trump Moves From Alternative Facts To Alternative Language“, *National Public Radio*, publicat în 2017, https://www.npr.org/2017/02/17/515630467/with-fake-news-trump-moves-from-alternative-facts-to-alternative-language.
67. S. Siddiqui, „Donald Trump faces backlash as he reveals «Fake News Awards» winners“, *The Guardian*, publicat în 2018, https://www.theguardian.com/us-news/2018/jan/17/trump-fake-news-awards-winners.
68. D. Phillips, „Brazil battles fake news «tsunami» amid polarized presidential election“, *The Guardian*, publicat în 2018, https://www.theguardian.com/world/2018/oct/10/brazil-fake-news-presidential-election-whatsapp-facebook.
69. C. Voinea, „We are stupid and contagious: infodemia și falsele ei remedii“, *Scena 9*, publicat online în 2020, https://www.scena9.ro/article/infodemie-dezinformare-coronavirus.
70. A.R. KhudaBukhsh *et al.*, „We Don’t Speak the Same Language: Interpreting Polarization through Machine Translation“, 2020. https://arxiv.org/abs/2010.02339.
71. Pew Research Center, „The shift in the American public’s political values: Political Polarization, 1994–2017, publicat în 2017, https://www.pewresearch.org/politics/interactives/political-polarization-1994-2017.
72. Pew Research Center, M. Jurkowitz *et al.*, „U.S. Media Polarization and the 2020 Election: A Nation Divided“, 2020, https://www.pewresearch.org/journalism/2020/01/24/u-s-media-polarization-and-the-2020-election-a-nation-divided.
72. Pew Research Center, „Dimok M, Wike R. America Is Exceptional in the Nature of Its Political Divide“, 2020, https://www.pewresearch.org/fact-tank/2020/11/13/america-is-exceptional-in-

the-nature-of-its-political-divide.
74. K. Hao, „How Facebook got addicted to spreading misinformation“, *MIT Technology Review*, publicat online în martie 2021, https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation; K. Hao, „The Facebook whistleblower says its algorithms are dangerous. Here’s why“, *MIT Technology Review*, publicat online în octombrie 2021, https://www.technologyreview. com/2021/ 10/05/1036519/facebook-whistleblower-frances-haugen-algorithms; D. Milmo, „How losing a friend to misinformation drove Facebook whistleblower“, *The Guardian*, publicat online în 2021, https://www. theguardian.com/technology/2021/oct/04/how-friend-lost-to-misinformation-drove-facebook-whistleblower-frances-haugen; G. Wells, J. Horwitz, D. Seetharaman, „Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show“, *The Wall Street Journal*, publicat la 14 septembrie 2021https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739, accesat la 6 noiembrie 2021.
75. E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You*, The Penguin Press, NY, 2011.
76. H. Allcott, M. Gentzkow, „Social media and fake news in the 2016 election“, *Journal of Economic Perspectives*, 2017, 31 (2), pp. 211–236.
77. M. Golebiewski, D. Boyd, „Data Voids. Where Missing Data Can Easily Be Exploited. Data & Society“, versiune actualizată în 2019, https://apo.org.au/sites/default/files/resource-files/2019-10/apo-nid265631.pdf.
78. „Sutherland Springs church shooting“, Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/wiki/Sutherland_ Springs_church_ shooting, accesat online la 28 ianuarie 2022.
79. L. Gornstein, „What is antifa? Is it a group or an idea, and what do supporters want?“, *CBS News*, publicat online la 29 martie 2021, https://www.cbsnews.com/news/what-is-antifa.
80. M. Fisher, J. Keller, „Why Does the U.S. Have So Many Mass Shootings? Research Is Clear: Guns“, *The New York Times*, publicat online la 7 noiembrie 2017, https://www.nytimes.com/2017/ 11/07/world/

americas/mass-shootings-us-international.html.
81. M.E. Hayden, „«Antifa» Responsible for Sutherland Springs Murders, According to Far-Right Media“, *Newsweek*, publicat online la 5 noiembrie 2017, https://www.newsweek.com/sutherland-springs-shooter-devin-patrick-kelley-was-antifa-according-far-right-702338.
82. „Stoneman Douglas High School shooting“, Wikipedia, The Free Encyclopedia, accesat online la 28 ianuarie 2022, https://en.wikipedia.org/wiki/Stoneman_Douglas_High_School_shooting.
83. R. Brandom, „One in five ad-serving websites is visited exclusively by fraud bots – The Verge“, *The Verge*, https://www.theverge. com/2017/5/24/15681080/ad-fraud-websites-traffic-bots-white-ops-report, accesat la 23 octombrie 2021.
84. A. Fou, „Ad «Relevancy» Is Fiction, And It’s Creepy“, *Forbes*, publicat online la 19 aprilie 2021, https://www.forbes.com/sites/augustinefou/2021/04/19/ad-relevancy-is-fiction-and-its-creepy, accesat la 23 octombrie 2021.
85. „Mobile Ad Fraud Schemes: How They Work, and How to Defend Against Them“, *Trend Micro*, publicat la 26 aprilie 2019, https://www.trendmicro.com/vinfo/us/security/news/mobile-safety/mobile-ad-fraud-schemes-how-they-work-and-how-to-defend-against-them, accesat la 23 octombrie 2021.
86. J.E.F. Teitcher *et al*., „Detecting, Preventing, and Responding to «Fraudsters» in Internet Research: Ethics and Tradeoffs“, *The Journal of law, medicine & ethics: a journal of the American Society of Law, Medicine & Ethics*. 2015, 43 (1), 116.
87. J.E.F. Teitcher *et al*., „Detecting, Preventing, and Responding to «Fraudsters» in Internet Research: Ethics and Tradeoffs“, *The Journal of law, medicine & ethics: a journal of the American Society of Law, Medicine & Ethics*. 2015, 43 (1), 116.
88. C. O'Sullivan, „What is Adversarial Machine Learning? An introduction to manipulating machine learning models“, *Medium*, publicat online la 12 iulie 2021, https://towardsdatascience.com/what-is-adversarial-machine-learning-dbe7110433d6.
89. I.L. Caragiale, „Bacalaureat“, *Wikisource*, publicat online la 3 iulie

2020, https://ro.wikisource.org/wiki/Bacalaureat.
90. B. Dickson, „Text-based AI models are vulnerable to paraphrasing attacks, researchers find“, *VentureBeat*, publicat online la 1 aprilie 2019, https://venturebeat.com/2019/04/01/text-based-ai-models-are-vulnerable-to-paraphrasing-attacks-researchers-find.
91. M.J., Coren, „All the things that still baffle self-driving cars, starting with seagulls“, *Quartz*, publicat online la 23 septembrie 2018, https://qz.com/1397504/all-the-things-that-still-baffle-self-driving-cars-starting-with-seagulls.
92. M. Himelein-Wachowiak, *et al*., „Bots and Misinformation Spread on Social Media: Implications for COVID-19“, *Journal of Medical Internet Research* 23.5 (2021), e26933.
93. T. Simonite, „Facebook Says Its New AI Can Identify More Problems Faster“, *Wired*, publicat online la 8 decembrie 2021, https://www.wired.com/story/facebook-says-new-ai-identify-more-problems-faster.
94. J. Hernandez, „A Military Drone With A Mind Of Its Own Was Used In Combat, U.N. Says“, *NPR*, publicat online la 1 iunie 2021, https://www.npr.org/2021/06/01/1002196245/a-u-n-report-suggests-libya-saw-the-first-battlefield-killing-by-an-autonomous-d.
95. T. Simonite, „It's Not Just the IRS – the US Government Wants Your Selfies“, *Wired*, publicat online la 28 ianuarie 2022, https://www.wired.com/story/irs-us-government-wants-selfies.
96. J. Dawes, „UN fails to agree on «killer robot» ban as nations pour billions into autonomous weapons research“, *The Conversation*, publicat online la 20 decembrie 2021, https://theconversation.com/un-fails-to-agree-on-killer-robot-ban-as-nations-pour-billions-into-autonomous-weapons-research-173616.
97. N. Al-Sibai, „MIT Professor Warns That Cartels Could Use «Slaughterbots» to Evade Justice“, *Futurism*, publicat online la 13 decembrie 2021, https://futurism.com/mit-professor-warns-cartels-could-use-slaughterbots.
98. E. Farge, „U.N. talks adjourn without deal to regulate «killer robots»“, *Reuters*, publicat online la 17 decembrie 2021, https://www.reuters.com/article/us-un-disarmament-idAFKBN2IW1UJ.

5. -În balanță: noi riscuri morale ale științei și societății algoritmice

1. M. O'Gieblyn, „My Music App Knows Me Way Too Well. Am I Stuck in a Groove?“, *Wired*, publicat online la 29 noiembrie 2021, https://www.wired.com/story/cloud-support-music-apps-predictability.
2. E. Rosenbaum, „IBM artificial intelligence can predict with 95% accuracy which workers are about to quit their jobs“, *CNBC*, publicat online la 3 aprilie 2019, https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html.
3. W. Coldwell, „What Happens When an AI Knows How You Feel?“, *Wired*, publicat online la 29 decembrie 2021, https://www.wired.com/story/artificial-emotional-intelligence.
4. B. Schillace, „Predicting Death Could Change the Value of a Life“, *Wired*, publicat online la 28 decembrie 2021, https://www.wired.com/story/predicting-death-value-life.
5. A. Završnik, „Criminal justice, artificial intelligence systems, and human rights“, *ERA Forum*, 2020, 20, pp. 567–583.
6. C.M. Drăgușin, „Stadiul actual al eticii în domeniul inteligenței artificiale“, *Dreptul*, 2020, 31 (3), pp. 90–102.
7. C.M. Drăgușin, „State of Wisconsin v. Loomis sau intersecția dintre inteligența artificială și dreptul fundamental la un proces echitabil“, *Dreptul*, 2020, 31 (4), pp.108–124.
8. D. McQuillan, „The Political Affinities of AI“, în A. Sudmann (ed.), *The Democratization of Artificial Intelligence*, Transcript-Verlag, Bielefeld, 2020, pp. 163–174, publicat online și accesat la 29 ianuarie 2022 la https://library.oapen.org/bitstream/handle/20.500.12657/43874/external_content.pdf.

6. REGULILE JOCULUI: SPRE UN MODEL EUROPEAN AL AUTONOMIEI UMANE

1. Y. Wang, „In China, the «Great Firewall» Is Changing a Generation“, *Politico*, publicat online în septembrie 2020, https://www.politico.com/news/magazine/2020/09/01/china-great-firewall-generation-405385.
2. „China issues rules on social credit system amid public concerns“, *Reuters*, publicat în 2020, https://www.reuters.com/world/china/china-issues-rules-social-credit-system-amid-public-concerns-2020-12-24,

accesat la 15 aprilie 2021.
3. J. Leibold, „Surveillance in China's Xinjiang Region: Ethnic Sorting, Coercion, and Inducement", *Journal of Contemporary China*, 2020, 29 (121), pp. 46–60; R. van Noorden, „The ethical questions that haunt facial-recognition research", *Nature*, 2020, 587, pp. 354–358, https://media.nature.com/original/magazine-assets/d41586-020-03187-3/d41586-020-03187-3.pdf.
4. European Commission, „Proposal for a Regulation of the European Parliament and of the Council, Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Ammending Certain Union Legislative Acts", publicat la 21 aprilie 2021, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206, accesat la 30 octombrie 2021.
5. R. Herschel, V.M. Miori, „Ethics & Big Data", *Technology in Society*, 2017, 49, pp. 31–36.
6. R. Herschel, V.M. Miori, „Ethics & Big Data", *Technology in Society*, 2017, 49, pp. 31–36.
7. R. Herschel, V.M. Miori, „Ethics & Big Data", *Technology in Society*, 2017, 49, pp. 31–36.
8. R. Herschel, V.M. Miori, „Ethics & Big Data", *Technology in Society*, 2017, 49, pp. 31–36.
9. R. Herschel, V.M. Miori, „Ethics & Big Data", *Technology in Society*, 2017, 49, pp. 31–36.
10. R. Herschel, V.M. Miori, „Ethics & Big Data", *Technology in Society*, 2017, 49, pp. 31–36.
11. J. Kokott, C. Sobotta, „The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR", *International Data Privacy Law*, 2013, 3 (4), pp. 222–228.
12. S. Rodotà, „Data Protection as a Fundamental Right", în *Reinventing Data Protection?*, Springer, Netherlands, 2009, pp. 77–82; Y. McDermott, „Conceptualising the right to data protection in an era of Big Data", *Big Data & Society*, 2017, 4 (1).
13. J. Kokott, C. Sobotta, „The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR",

International Data Privacy Law, 2013, 3 (4), pp. 222–228.
14. J. Kokott, C. Sobotta, „The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR“, *International Data Privacy Law*, 2013, 3 (4), pp. 222–228.
15. J. Kokott, C. Sobotta, „The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR“, *International Data Privacy Law*, 2013, 3 (4), pp. 222–228.
16. S. Rodotà, „Data Protection as a Fundamental Right“, în *Reinventing Data Protection?*, Springer, Netherlands, 2009, pp. 77–82.
17. S. Rodotà, „Data Protection as a Fundamental Right“, în *Reinventing Data Protection?*, Springer, Netherlands, 2009, pp. 77–82.
18. Y. McDermott, „Conceptualising the right to data protection in an era of Big Data“, *Big Data & Society*, 2017, 4 (1).
19. S. Rodotà, „Data Protection as a Fundamental Right“, în *Reinventing Data Protection?*, Springer, Netherlands, 2009, pp. 77–82.
20. S. Rodotà, „Data Protection as a Fundamental Right“, în *Reinventing Data Protection?*, Springer, Netherlands, 2009, pp. 77–82.
21. E.R.Tufte, „Computer Literacy Bookshops Interview, 1994-1997“, publicat în 1997, https://www.edwardtufte.com/tufte/complit_9497, accesat la 31 octombrie 2021.
22. J. Orlowki, „The Social Dilemma“, Netflix, 2020, https://www.netflix.com/title/81254224.
23. Norwegian Consumer Council, „Out of Control. How Consumers Are Exploited by the Online Advertising Industry“; 2020; J. Duportail, „I asked Tinder for my data. It sent me 800 pages of my deepest, darkest secrets“, *The Guardian*, publicat la 26 septembrie 2016, https:// www.theguardian.com/technology/2017/sep/26/tinder-personal-data-dating-app-messages-hacked-sold.
24. Redacţia The Pillar, „Pillar Investigates: USCCB gen sec Burrill resigns after sexual misconduct allegations“, *The Pillar*, publicat online la 20 iulie 2021, https://www.pillarcatholic.com/p/pillar-investigates-usccb-gen-sec, accesat la 31 octombrie 2021.
25. World Economic Forum, „Measuring Stakeholder Capitalism Towards Common Metrics and Consistent Reporting of Sustainable Value

Creation“, publicat online în septembrie 2020, https://www3.weforum.org/docs/WEF_IBC_Measuring_Stakeholder_Capitalism_Report_2020.pdf, accesat la 31 octombrie 2021.
26. K. Hao, „How Facebook got addicted to spreading misinformation“, *MIT Technology Review*, publicat online în martie 2021, https://www.technologyreview.com/2021/03/11/1020600/facebook-responsible-ai-misinformation.
27. E.M. Bender *et al.*, „On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?“, în *FAccT ’21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ACM, 2021.
28. J. Carrie, „More than 1,200 Google workers condemn firing of AI scientist Timnit Gebru“, *The Guardian*, publicat online la 4 decembrie 2020, https://www.theguardian.com/technology/2020/dec/04/timnit-gebru-google-ai-fired-diversity-ethics, accesat la 31 octombrie 2021.
29. J. Buolamwini, T. Gebru, „Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification“, în *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 2018, http://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline, accesat la 31 octombrie 2021.
30. J. Vincent, „Google is Poisoning its Reputation with AI Researchers“, *The Verge*, publicat online la 13 aprilie 2021, https://www. theverge.com/2021/4/13/22370158/google-ai-ethics-timnit-gebru-margaret-mitchell-firing-reputation, accesat la 31 octombrie 2021.
31. „Inside the Tobacco Deal: Jeffrey Wigand“, *PBS*, publicat în 2014, https://www.pbs.org/wgbh/pages/frontline/shows/settlement/timelines/wigand.html, accesat la 31 octombrie 2021.
32. E. Lopatto, „How Do You Cross-Examine a Whistleblower?“, *The Verge*, publicat online la 15 septembrie 2021, https://www.theverge.com/2021/9/15/22676789/elizabeth-holmes-theranos-wire-fraud-erika-cheung-whistleblower, accesat la 31 octombrie 2021.
33. „The Cambridge Analytica Files“, *The Guardian*, publicat online în 2018, https://www.theguardian.com/news/series/cambridge-analytica-files, accesat la 31 octombrie 2021.

34. E. Pilkington, „Google’s secret cache of medical data includes names and full details of millions – whistleblower“, *The Guardian*, publicat online la 12 noiembrie 2019, https://www.theguardian.com/technology/2019/nov/12/google-medical-data-project-nightingale-secret-transfer-us-health-information, accesat la 31 octombrie 2021.
35. K. Hao, „The Facebook whistleblower says its algorithms are dangerous. Here’s why“, *MIT Technology Review*, publicat online în octombrie 2021, https://www.technologyreview.com/2021/10/05/1036519/ facebook-whistleblower-frances-haugen-algorithms.
36. O. Keynes, „The Facebook Whistleblower Won’t Change Anything“, *Wired*, publicat online la 18 octombrie 2021, https://www.wired.com/story/facebook-whistleblower-wont-change-anything, accesat la 31 octombrie 2021.
37. La Quadrature du Net, „Amazon fined 746 million euros following our collective legal action“, publicat online în 2021, https://www.laquadrature.net/en/2021/07/30/amazon-fined-746-million-euros-following-our-collective-legal-action.
38. La Quadrature du Net, „First sanction against Google following our collective complaints“, publicat online în 2019, https://www.laquadrature.net/en/2019/01/21/first-sanction-against-google-following-our-collective-complaints.
39. M. Burgess, „Europe’s Move Against Google Analytics Is Just the Beginning“, *Wired*, publicat online în 2022, https://www.wired.com/story/google-analytics-europe-austria-privacy-shield.
40. N. Lomas, „In bad news for US cloud services, Austrian website’s use of Google Analytics found to breach GDPR“, *TechCrunch*, publicat online în 2022, https://techcrunch.com/2022/01/12/austrian-dpa-schrems-ii.
41. O. Keynes, „The Facebook Whistleblower Won’t Change Anything“, *Wired*, publicat online la 18 octombrie 2021, https://www.wired.com/story/facebook-whistleblower-wont-change-anything, accesat la 31 octombrie 2021.

42. C. Rughiniș, M.G. Flaherty, „Algorithmic discrimination and its temporal logics“, în *Digital Humanities in the Nordic Countries*, DHN, 2019, https://cst.dk/DHN2019Pro/abstracts/Rughinis-Flaherty-DHN%20abstract-FINAL.pdf, accesat la 8 noiembrie 2021.
43. K. Johnson, „Facebook Drops Facial Recognition to Tag People in Photos“, *Wired*, publicat online la 2 noiembrie 2021, https://www.wired.com/story/facebook-drops-facial-recognition-tag-people-photos.
44. L. Matsakis, „Amazon Won’t Let Police Use Its Facial-Recognition Tech for One Year“, *Wired*, publicat online pe 10 iunie 2020, https://www.wired.com/story/amazon-facial-recognition-police-one-year-ban-rekognition.
45. T. Simonite, „Face Recognition Is Being Banned – but It’s Still Everywhere“, *Wired*, publicat online la 22 decembrie 2021, https://www.wired.com/story/face-recognition-banned-but-everywhere.
46. T. Simonite, „It's Not Just the IRS – the US Government Wants Your Selfies“, *Wired*, publicat online la 28 ianuarie 2022, https://www.wired.com/story/irs-us-government-wants-selfies.
47. European Parliament, „European Parliament resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters“, Publicat 6 octombrie 2021.
48. R. Rughiniș, C. Rughiniș, S.N. Vulpe, D. Rosner, „From social netizens to data citizens: variations of GDPR awareness in 28 European countries“, *Computer Law & Security Review*, publicat online în septembrie 2021, https://www.sciencedirect.com/science/article/abs/pii/S02673 64921000583, accesat la 8 noiembrie 2021.
49. R. McKean, E. Kurowska-Tober, H. Waem, „DLA Piper GDPR fines and data breach survey: January 2021“, *DLA Piper Publications*, publicat online la 19 ianuarie 2021, https://www.dlapiper.com/en/uk/insights/publications/2021/01/dla-piper-gdpr-fines-and-data-breach-survey-2021, accesat la 6 noiembrie 2021.
50. M. Burgess, „Amazon’s Massive GDPR Fine Shows the Law’s Power – and Limits“, *Wired*, publicat online la 8 mai 2021, https://www.wired.com/story/amazon-gdpr-fine-shows-power-limits, accesat la 6 noiembrie 2021.

51. V. Manancourt, M. Scott, „What the US can teach Europe about privacy“, *Politico*, publicat online în 2021, https://www.politico.eu/article/europe-privacy-fines-tech-what-the-us-can-teach, accesat la 7 noiembrie 2021.
52. J. Taylor, „Big Tech Turns Its Lobbyists Loose on Europe, Alarming Regulators“, *The New York Times*, publicat online la 14 decembrie 2020, https://www.nytimes.com/2020/12/14/technology/big-tech-lobbying-europe.html, accesat la 7 noiembrie 2021.
53. J. Keane, „Italy has quietly become one of Big Tech’s most prolific antagonists“, *CNBC*, publicat online la 18 ianuarie 2022, https://www.cnbc.com/2022/01/18/european-regulators-in-italy-france-germany-and-uk-rein-in-big-tech-.html.
54. „Europe's battle to curb Big Tech“, France24, publicat online la 10 noiembrie 2021, https://www.france24.com/en/live-news/20211110-europe-s-battle-to-curb-big-tech.
55. M. Armstrong, „Google's Contested EU Antitrust Fine is Peanuts“, *Statista*, publicată online la 12 februarie 2020, https://www.statista.com/chart/20805/google-antitrust-fine-time-to-earn.
56. A. Yanchur *et al*., „Computer says No: How the EU’s AI laws cause new injustice“, *EUobserver*, publicat online la 23 august 2021, https://euobserver.com/investigations/152695, accesat la 31 octombrie 2021.
57. T. Simonite, „This Group Pushed More AI in US Security – and Boosted Big Tech“, *Wired*, publicat online la 1 noiembrie 2021, https://www.wired.com/story/group-pushed-ai-us-security-boosted-tech, accesat la 7 noiembrie 2021.
58. T. Simonite, „Congress Is Eyeing Face Recognition, and Companies Want a Say“, *Wired*, publicat online la 23 noiembrie 2020, https://www.wired.com/story/congress-eyeing-face-recognition-companies-want-say, accesat la 8 noiembrie 2021.
59. T. Klosowski, „The State of Consumer Data Privacy Laws in the US (And Why It Matters)“, *The New York Times*, publicat la 6 septembrie 2021, https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us, accesat la 7 noiembrie 2021.
60. R. McKean, E. Kurowska-Tober, H. Waem, „DLA Piper GDPR fines

and data breach survey: January 2021“, *DLA Piper Publications*, publicat online la 19 ianuarie 2021, https://www.dlapiper.com/en/uk/insights/publications/2021/01/dla-piper-gdpr-fines-and-data-breach-survey-2021, accesat la 6 noiembrie 2021.
61. T. Wertheimer, „Blake Lemoine: Google fires engineer who said AI tech has feelings“, BBC News, publicat online la 23 iulie 2022, https:// www.bbc.com/news/technology-62275326, accesat la 1 septembrie 2022.
62. B. Lemoine, „Is LaMDA Sentient? – an Interview“, *Medium*, publicat online la 11 iunie 2022, https://cajundiscordian.medium.com/is-lamda-sentient-an-interview-ea64d916d917, accesat la 1 septembrie 2022.
63. E.M. Bender *et al*., „On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?“, *FAccT 21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, martie 2021, pp. 610–623, disponibil online la https://doi.org/10.1145/ 3442188. 3445922.
64. C. Stross, „Artificial Intelligence: Threat or Menace?“, blogul Charlie’s Diary, publicat online în decembrie 2019, https://www. antipope. org/charlie/blog-static/2019/12/artificial-intelligence-threat.html, accesat la 1 septembrie 2022.
65. J. Kantor *et al*., „The Rise of the Worker Productivity Score“, *The New York Times*, publicat online la 14 august 2022, https://www.nytimes. com/interactive/2022/08/14/business/worker-productivity-tracking. html, accesat la 1 septembrie 2022.
66. C. Doctorow, „Workplace surveillance is coming for you“, *Medium*, publicat online la 21 august 2022, https://doctorow.medium.com/ workplace-surveillance-is-coming-for-you-8720f8d8109f, accesat la 1 septembrie 2022.
67. Center for Reproductive Rights, „After Roe Fell: Abortion Rights by State“, https://reproductiverights.org/maps/abortion-laws-by-state, accesat la 22 septembrie 2022.
68. D. Cameron, „Senator: Overturning Roe Will «Weaponize» Women’s Data“, *Gizmondo*, publicat online la 3 martie, 2022https://gizmodo. com/roe-v-wade-supreme-court-abortion-overturn-ron-wyden-da-1848876732, accesat la 1 septembrie 2022.

69. J. Cox, „Location Data Broker Admits Users Accessed Data for Planned Parenthood Clinics“, *Vice*, publicat online la 8 iulie 2022, https://www.vice.com/en/article/xgyxd7/location-data-broker-placer-admits-accessed-planned-parenthood-data, accesat la 1 septembrie 2022.
70. J. Cox, „Data Broker Is Selling Location Data of People Who Visit Abortion Clinics“, Vice, Publicat online la 3 mai 2022, https://www.vice.com/en/article/m7vzjb/location-data-abortion-clinics-safegraph-planned-parenthood, accesat la 1 septembrie 2022.
71. J. Cox, „Over 40 Members of Congress Urge Google to Limit Location Data Collection for a Post-Roe America“, *Vice*, publicat online la 24 mai 2022, https://www.vice.com/en/article/v7d5m3/40-members-of-congress-google-stop-location-data-abortion-rights, accesat la 1 septembrie 2022.
72. N. Singer, B.X. Chen, „In a Post-Roe World, the Future of Digital Privacy Looks Even Grimmer“, *The New York Times*, Publicat online la 13 iulie 2022, https://www.nytimes.com/2022/07/13/technology/personaltech/abortion-privacy-roe-surveillance.html, Accesat la 22 septembrie 2022.
73. S. Geoghegan, L. Fellow, D. Khabbaz, „Reproductive Privacy in the Age of Surveillance Capitalism“, *EPIC*, publicat online la 7 iulie 2022, https://epic.org/reproductive-privacy-in-the-age-of-surveillance-capitalism, accesat la 1 septembrie 2022.
74. D. Cameron, „Senator: Overturning Roe Will «Weaponize» Women's Data“, *Gizmondo*, publicat online la 3 martie, 2022https://gizmodo.com/roe-v-wade-supreme-court-abortion-overturn-ron-wyden-da-1848876732, accesat la 1 septembrie 2022.
75. C. Corcoran, „Ukraine credits Turkish drones with eviscerating Russian tanks and armor in their first use in a major conflict“, *Business Insider*, publicat online la 28 februarie 2022, https://www.businessinsider.com/ukraine-hypes-bayraktar-drone-as-videos-show-destroyed-russia-tanks-2022-2, accesat la 1 septembrie 2022.
76. Reuters, „Zelenskiy says Turkish drone maker to build Ukraine factory“, publicat online la 9 septembrie 2022, https://www.reuters.com/business/aerospace-defense/zelenskiy-says-turkish-drone-maker-

build-ukraine-factory-2022-09-09, accesat la 15 septembrie 2022.
77. L. Hudson, P. McLeary, A. Warad, „U.S. sending new drones to Ukraine ahead of southern push“, *Politico*, publicat online la 19 august 2022, politico.com/newsletters/national-security-daily/2022/08/19/u-s-sending-new-drones-to-ukraine-ahead-of-southern-push-00052856#, accesat la 1 septembrie 2022.
78. „Norwegian-developed drone to Ukraine“, publicat online la 24 august 2022, https://www.regjeringen.no/en/aktuelt/droner/id2924942, accesat la 1 septembrie 2022.
79. I. Kossov, „A Game of Drones: Ukraine builds up UAV fleet“, *The Kyiv Independent*, publicat online la 26 iulie 2022, https://kyivindependent.com/national/a-game-of-drones-ukraine-builds-up-uav-fleet, accesat la 1 septembrie 2022.
80. A.E. Kramer, „Russia Deploys Iranian Attack Drones, Ukrainian Military Says“, *The New York Times*, publicat online la 18 septembrie 2022, https://www.nytimes.com/2022/09/18/world/europe/iran-attack-drones-ukraine.html, accesat la 22 septembrie 2022.
81. C. Vallance, „Ukraine sent dozens of «dronations» to build army of drones“, BBC News, publicat online la 8 iulie 2022, https://www.bbc.com/news/technology-62048403, accesat la 1 septembrie 2022.
82. I. Kossov, „A Game of Drones: Ukraine builds up UAV fleet“, *Kyiv Independent*, publicat online la 26 iulie 2022, https:// kyivindependent.com/national/a-game-of-drones-ukraine-builds-up-uav-fleet, accesat la 1 septembrie 2022.
83. „Ukraine conflict: How are drones being used?“, BBC News, publicat online la 22 august 2022, https://www.bbc.com/news/world-62225830, accesat online la 1 septembrie 2022.
84. J. Clayton, „How facial recognition is identifying the dead in Ukraine“, BBC News, publicat online la 13 aprilie 2022, https://www. bbc.com/news/technology-61055319, accesat la 1 septembrie 2022.
85. P. Dave, J. Dastin, „Exclusive: Ukraine has started using Clearview AI’s facial recognition during war“, Reuters, publicat online la 14 martie 2022, https://www.reuters.com/technology/exclusive-ukraine-has-started-using-clearview-ais-facial-recognition-during-war-2022-03-13, accesat la

1 septembrie 2022.
86. „Clearview AI“, https://en.wikipedia.org/wiki/Clearview_AI, actualizat online la 19 septembrie 2022, accesat la 22 septembrie 2022.
87. K. Hill, „The Secretive Company That Might End Privacy as We Know It“, *The New York Times*, publicat online la 18 ianuarie 2020, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html, accesat la 10 ianuarie 2021.
88. P. Tucker, „US Working on AI to Predict Ukraine’s Ammo and Weapons Needs“, *Defense One*, publicat online la 20 septembrie 2022, https://www.defenseone.com/technology/2022/09/us-working-ai-predict-ukraines-ammo-and-weapons-needs/377429, accesat la 22 septembrie 2022.
89. A. Shamoug, „In Ukraine, machine-learning algorithms and big data scans used to identify war-damaged infrastructure“, UNDP, publicat online la 5 iulie 2022, https://www.undp.org/blog/ukraine-machine-learning-algorithms-and-big-data-scans-used-identify-war-damaged-infrastructure, accesat la 1 septembrie 2022.
90. N. Werkhäuser, „«Killer robots»: Will they be banned?“, Deutsche Welle, publicat online la 25 iulie 2022, https://www.dw.com/en/killer-robots-will-they-be-banned/a-62587436, accesat la 1 septembrie 2022.
91. M. Heikkilä, „Why business is booming for military AI startups“, *MIT Technology Review*, publicat online la 7 iulie 2022, https://www.technologyreview.com/2022/07/07/1055526/why-business-is-booming-for-military-ai-startups, accesat la 1 septembrie 2022.
92. M. Whittaker, L. Suchman, „The Myth of Artificial Intelligence“, *The American Prospect*, publicat online la 8 decembrie 2021, https://prospect.org/culture/books/myth-of-artificial-intelligence-kissinger-schmidt-huttenlocher, accesat la 1 septembrie 2022.
93. A. Holland, „Known Unknowns: Data Issues and Military Autonomous Systems“, UNIDIR, publicat online la 17 mai 2021, https://unidir.org/known-unknowns, accesat la 1 septembrie 2022.
94. N. Lomas, „Oracle’s «surveillance machine» targeted in US privacy class action“, *Tech Crunch*, publicat online la 22 august 2022, https://techcrunch.com/2022/08/22/oracle-us-privacy-class-action,

accesat la 22 septembrie 2022.
95. P. Dave, D. Bartz, „California alleges Amazon stifled price competition in lawsuit“, Reuters, publicat online la 15 septembrie 2022, https://www.reuters.com/legal/california-files-lawsuit-against-amazoncom-allegedly-blocking-price-competition-2022-09-14, accesat la 22 septembrie 2022.
96. C. Doctorow, „California’s antitrust case against Amazon“, *Medium*, publicat online la 15 septembrie 2022, https://doctorow. medium. com/californias-antitrust-case-against-amazon-82b8e4177fb7, accesat la 22 septembrie 2022.
97. Comisia Europeană, „The Digital Services Act package“, publicat online la 5 iulie 2022, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package, accesat la 1 septembrie 2022.
98. J. Espinoza, „How Big Tech lost the antitrust battle with Europe“, *Ars Tehnica*, publicat online la 21 martie 2022, https://arstechnica. com/tech-policy/2022/03/how-big-tech-lost-the-antitrust-battle-with-europe, accesat la 1 septembrie 2022.
99. P.F. Drucker, *Landmarks of Tomorrow: A Report on the New „Post-Modern“ World*, Transaction Publishers, London, 1957.
100. F. Weinert, *Copernicus, Darwin, & Freud: Revolutions in the History and Philosophy of Science*, Wiley-Blackwell, West Sussex, 2009.

În colecţia **Ştiinţă** au mai apărut

Marcus Chown, *Magicienii*
Brian Greene, *Până la sfârşitul timpului*
Mircea Flonta, *Folozofia cercetătorului: Înţelegerea fizicii cuantice la Niels Bohr şi Werner Heisenberg*
Richard Dawkins, Yan Wong, *Povestea strămoşilor noştri: Pelerinaj la începuturile vieţii*
Mihai G. Netea, *O istorie genetică (incompletă) a românilor*
Alexandru N. Stermin, *Căzuţi din junglă: Poveştile unui explorator*
Mario Livio, *De ce?: Ce ne stârneşte curiozitatea*
Sverker Johansson, *Zorii limbajului*
Gaia Vince, *Focul, limbajul, frumusețea, timpul: Cum au depășit oamenii legile firești ale evoluției*
Jim Al-Khalili, *Ce ne spune fizica despre lume*
Michael White, John Gribbin, *Stephen Hawking: O viață dedicată științei*
Robert Levine, *Geografia timpului: De ce timpul nu se măsoară la fel în toate culturile*
Carlo Rovelli, *Și dacă timpul n-ar exista?*
Antonio Damasio, *Strania ordine a lucrurilor: Viața, sentimentele și nașterea culturilor*
Pier Vincenzo Piazza, *Homo biologicus: Cum explică biologia natura umană*
Paul Davies, *Rețelele de informație și misterul vieții*
John D. Barrow, *1 + 1 nu face (întotdeauna) 2: O lecție de matematică*
Paul Nurse, *Ce este viața? Cinci mari idei ale biologiei*
Noreena Hertz, *Secolul singurătății: O pledoarie pentru relațiile interumane*
Frank Wilczek, *Fundamentele lumii fizice*
Alexandru Babeş, *Povestea creierului: În căutarea celui mai complicat obiect din univers*
Jim Baggott, *Masa: Încercarea de a înțelege materia, de la atomii grecilor la câmpurile cuantice*
Alexandru N. Stermin, *Călătorie în jurul omului*
William H. Calvin, *O scurtă istorie a minții: De la maimuțele antropoide la*

intelect și mai departe

Marco Andreatta, *Forma lucrurilor: Alfabetul geometriei*

Jim Al-Khalili, *Găuri negre, găuri de vierme și călătoria în timp*

Alexis Willett, Jennifer Barnett, *De cât de mult creier avem cu adevărat nevoie? Metoda infailibilă ca să-ți folosești 100% inteligența*

David Whitehouse, *Călătorie spre centrul Pământului: Extraordinara expediție de explorare științifică în adâncurile lumii noastre*

Carlo Rovelli, *Helgoland: Cum să înțelegem teoria cuantică*

Michael Graziano, *Conștiința dintr-o nouă perspectivă: O teorie științifică a experienței subiective*

Stephen Hawking, *Răspunsuri scurte la marile întrebări*

Adrian Bejan, *Fizica vieții: Evoluția pretutindeni*

David Eagleman, Anthony Brandt, *Specia rebelă: Despre creativitatea oamenilor și despre modul în care ea schimbă lumea*

Frans de Waal, *Mama. Ultima îmbrățișare: Ce ne spun emoțiile animalelor despre noi înșine*

Jim Al-Khalili (ed.), *Extratereștrii: Ce ne spune știința despre viața în univers*

Alexandru Toma Pătrașcu, *Povestiri despre epidemii și vaccinuri*

Oliver Sacks, *Fluviul conștiinței*

Sabine Hossenfelder, *Rătăciți printre formule: Cum îi derutează frumusețea pe fizicieni*

Hélène Courtois, *Călătorie pe valuri de galaxii: Laniakea, noua noastră adresă în Univers*

Silvana Condemi, François Savatier, *Sapiens: Cele mai recente descoperiri*

Bill Mesler, H. James Cleaves II, *Scurtă istorie a creației: Știința și căutarea originii vieții*

Manfred Spitzer, *Demența digitală: Cum ne tulbură mintea noile tehnologii*

Richard Muller, *Acum: Fizica timpului*

John Hands, *Cosmosapiens: Evoluția omului de la originile universului*

Alex Doppelgänger, *Pământul este plat! Cum să ne ferim de capcanele pseudoștiinței*

Pietro Greco, *Povestea numărului Π*

Scott Stossel, *Anxietatea: O poveste personală despre frică, speranță și*

căutarea liniștii interioare

Frans de Waal, *Suntem îndeajuns de inteligenţi pentru a înţelege inteligența animalelor?*

Stephen Hawking, *Găurile negre: Prelegerile Reith, BBC*

Carlo Rovelli, *Realitatea nu e ceea ce pare: Structura fundamentală a lucrurilor*

Max Tegmark, *Viața 3.0: Omul în epoca inteligenței artificiale*

Carlo Rovelli, *Ordinea timpului*

Jean-Michel Besnier, Laurent Alexandre, *Pot face roboţii dragoste? 12 întrebări despre transumanism*

Edward O. Wilson, *Originile creativităţii umane*

Norbert Wiener, *Dumnezeu şi Golemul: Comentariu asupra câtorva probleme în care cibernetica intră în contradicţie cu religia*

Richard Dawkins, *Candelă de-o clipă-n întuneric: Viaţa mea în lumea științei*

Christine Kenneally, *Povestea secretă a speciei umane: Cum ne sunt modelate identitatea și viitorul de ADN și de istorie*

Paolo Zellini, *Matematica zeilor și algoritmii oamenilor*

Jim Baggott, *Origini: Povestea științifică a creației*

Lewis Dartnell, *Cunoașterea: Cum poate fi reclădită civilizația după un cataclism*

Jacob Bronowski, *Ascensiunea omului*

David Eagleman, *Creierul: Povestea noastră*

Kristian Marlow, Berit Brogaard, *Mintea supraumană: Sau despre genialitate ca însușire firească a creierului nostru*

Edward O. Wilson, *Sensul existenței umane*

Oliver Sacks, *Unchiul Tungsten: Amintiri dintr-o copilărie chimică*

Frans de Waal, *Bonobo și ateul: În căutarea umanismului printre primate*

Amir Alexander, *Infinitezimal: Cum a contribuit la făurirea lumii moderne o teorie matematică periculoasă*

Steven Weinberg, *Lumea explicată: Descoperirea științei moderne*

Vincent Deary, *Cum suntem: O psihologie a rutinei*

Alice Roberts, *Marea șansă a existenței: Evoluția și dezvoltarea omului*

Richard Dawkins, *Lumea ca un mare spectacol: Dovezile evoluției*

Carlo Rovelli, *Șapte scurte lecții de fizică*

Mario Livio, *Gafe strălucite: Greşeli colosale ale unor mari savanţi, care ne-au schimbat perspectiva asupra vieţii şi universului*

Freeman Dyson, *Originile vieții*

Oliver Sacks, *Halucinații*

Richard P. Feynman, *Sensul tuturor lucrurilor: Gândurile unui cetățean și om de știință*

David Eagleman, *Incognito: Viețile secrete ale creierului*

Antonio Damasio, *Sinele: Construirea creierului conştient*

Martin Bojowald, *Ce a fost înainte de Big Bang? O istorie întreagă a universului*

Douglas R. Hofstadter, *Gödel, Escher, Bach: Brilianta Ghirlandă Eternă*

Jim Baggott, *Higgs: Inventarea și descoperirea „Particulei lui Dumnezeu“*

Thomas Metzinger, *Tunelul Eului: Știința minții și mitul sinelui*

Stephen Hawking, *Scurtă istorie a vieții mele*

Richard Dawkins, *O curiozitate fără margini: Cum am devenit om de știință*

Ian Stewart, *Dă oare Dumnezeu cu zarul? Noua matematică a haosului*

Michael Shermer, *De ce e Darwin important*

Paul J. Zak, *Molecula morală: Sursa iubirii şi a prosperităţii*

Oliver Sacks, *Ochiul minţii: Cum se vede lumea când vederea dă greş*

Stephen Hawking, *Teoria universală: Originea şi soarta Universului*

Andrew Solomon, *Demonul amiezii: O anatomie a depresiei*

Cédric Villani, *Teorema vie*

Edward O. Wilson, *Cucerirea socială a pământului*

Richard P. Feynman, *Vă ţineţi de glume, domnule Feynman! Aventurile unui personaj ciudat*

Leonard Mlodinow, *Subliminal: Cum ne determină inconştientul comportamentul*

Lawrence M. Krauss, *Omul cuantic: Biografia ştiinţifică a lui Richard Feynman*

Niels Bohr, *Teoria atomică şi descrierea naturii: Patru eseuri şi un studiu introductiv*

Oliver Sacks, *Un picior de sprijin*

Simon Singh, Edzard Ernst, *Adevărul despre medicina alternativă*

David Berlinski, *Unu, doi, trei: Matematica absolut elementară*
Oliver Sacks, *Văzând glasuri: O călătorie în lumea surzilor*
Daniel M. Wegner, *Iluzia voinţei conştiente*
Stephen Hawking, Leonard Mlodinow, *Marele plan*
Leonard Susskind, *Peisajul cosmic: Teoria corzilor şi iluzia unui plan inteligent*
Richard Dawkins, *Magia realităţii: De unde ştim care este de fapt adevărul*
Oliver Sacks, *Revenirea la viaţă*
Mario Livio, *Este Dumnezeu matematician?*
Ian Stewart, *Îmblânzirea infinitului: Povestea matematicii*
Roger Scruton, *Cultura modernă pe înţelesul oamenilor inteligenţi*
Keith Devlin, *Partida neterminată: Pascal, Fermat şi scrisoarea din secolul al XVII-lea care a creat lumea modernă*
Daniel J. Levitin, *Creierul nostru musical: Ştiinţa unei eterne obsesii*
Ian Stewart, *De ce frumuseţea este adevărul: O istorie a simetriei*
Antonio Damasio, *În căutarea lui Spinoza: Cum explică ştiinţa sentimentele*
Oliver Sacks, *Muzicofilia: Povestiri despre muzică și creier*
Richard Dawkins, *Ceasornicarul orb*
Atul Gawande, *Pe muchie de cuțit: Însemnările unui chirurg despre o ştiinţă imperfectă*
Alessandro Baricco, *Barbarii: Eseu despre mutație*
Brian Greene, *Universul elegant: Supercorzi, dimensiuni ascunse și căutarea teoriei ultime*
John D. Barrow, *Cartea infinitului: Scurtă introducere în nemărginit, etern şi nesfârşit*
Simon Singh, *Big Bang: Originea universului*
Mario Livio, *Ecuația care n-a putut fi rezolvată: Matematicieni de geniu rezolvă limbajul simetriilor*
Jeremy Bernstein, *Albert Einstein și frontierele fizicii*
Steven Weinberg, *Visul unei teorii finale*
Stephen Hawking, Leonard Mlodinow, *O mai scurtă istorie a timpului*
Richard P. Feynman, *Șase lecții ușoare: Bazele fizicii explicate de cel mai strălucit profesor*
Charles Seife, *Zero: Biografia unei idei periculoase*

Clifford A. Pickover, *Banda lui Mobius*
Steven Weinberg, *Descoperirea particulelor subatomice*
Georges Charpak, Henri Broch, *Lecţii de vrăjitorie: Ştiinţa şi paranormalul*
Oliver Sacks, *Omul care îşi confunda soţia cu o pălărie*
Ernst Mayr, *De la bacterii la om: Evoluţia lumii vii*
Ian Stewart, *Numerele naturii: Ireala realitate a imaginaţiei matematice*
Paul Davies, *Ultimele trei minute*
Albert Einstein, *Cum văd eu lumea: O antologie*
Albert Einstein, *Teoria relativităţii pe înţelesul tuturor*